

RIGOROUS DERIVATION OF NONLINEAR SCALAR CONSERVATION LAWS FROM FOLLOW-THE-LEADER TYPE MODELS VIA MANY PARTICLE LIMIT

M. DI FRANCESCO AND M.D. ROSINI

ABSTRACT. We prove that the unique entropy solution to a scalar nonlinear conservation law with *strictly monotone velocity* and nonnegative initial condition can be rigorously obtained as the large particle limit of a microscopic follow-the-leader type model, which is interpreted as the discrete Lagrangian approximation of the nonlinear scalar conservation law. More precisely, we prove that the empirical measure (respectively the discretised density) obtained from the follow-the-leader system converges in the 1-Wasserstein topology (respectively in $\mathbf{L}_{\text{loc}}^1$) to the unique Kruzkov entropy solution of the conservation law. The initial data are taken in \mathbf{L}^∞ , nonnegative, and with compact support, hence we are able to handle densities with vacuum. Our result holds for a reasonably general class of velocity maps (including all the relevant examples in the applications, e.g. in the Lighthill-Whitham-Richards model for traffic flow) with possible degenerate slope near the vacuum state. The proof of the result is based on discrete \mathbf{BV} estimates and on a discrete version of the one-sided Oleinik-type condition. In particular, we prove that the regularizing effect $\mathbf{L}^\infty \mapsto \mathbf{BV}$ for nonlinear scalar conservation laws is intrinsic of the discrete model.

Keywords: Micro-macro limit and Scalar conservation laws and Follow-the-leader models and Oleinik condition and Entropy solutions and Particle method.

2010 AMS Subject classification: 35L65 and 35L45 and 90B20 and 65N75 and 82C22.

1. INTRODUCTION

1.1. Nonlinear scalar conservation laws in one space dimension. A scalar conservation law in one space dimension for the unknown variable ρ is a first order Partial Differential Equation (PDE) of the form

$$\rho_t + f(\rho)_x = 0, \quad t > 0, \quad x \in \mathbb{R}, \quad (1.1)$$

see e.g. the books [11, 18, 49] as general references. The unknown ρ typically describes the *mass density* of a given medium with a one-dimensional structure. In this case, a conservation law of the form (1.1) can be derived under the hypothesis that the time evolution of the (linear) mass contained in an arbitrary interval $]a, b[\subset \mathbb{R}$ at time $t > 0$, namely the integral $\int_a^b \rho(t, x) dx$, is only affected by the *flux* $f(\rho)$ of the medium at the edges of the interval, $x = a$ and $x = b$. In this sense, (1.1) can be seen as a *continuity equation* for ρ , and the expression

$$f(\rho) = \rho v(\rho) \quad (1.2)$$

arises very naturally, where $\rho \mapsto v(\rho)$ is a constitutive law for the *Eulerian velocity* of the medium.

The formulation of the equation (1.1) relies on the *continuum assumption*, in which the medium is assumed to be indefinitely divisible without changing its physical nature. Such assumption is needed to introduce the concept of macroscopic local density ρ , defined as the limit of the ratio $\Delta\ell/\Delta x$ as the measure of the elementary interval Δx goes to zero, being $\Delta\ell$ the (linear) mass contained in Δx . The continuum assumption is ensured by considering a *very large mass* compared to the size of the domain occupied by the medium.

The density of vehicles in traffic flow is a typical example from real world applications in which an equation of the form (1.1) is used as the ‘macroscopic counterpart’ of a large system of moving vehicles. Traffic flow is indeed one of the main motivating applications behind the present paper, and we shall therefore describe the use of the continuum assumption in this context in Section 1.5.

The nonlinearity in the constitutive law $f = f(\rho)$ of the flux has made the mathematical theory for (1.1) a challenging topic through the past decades. In Section 2.1 we recall the basic notions of the classical *theory of entropy solutions* for scalar conservation laws, the foundations of which go back to the pioneering papers by Oleinik [39] and Kruzkov [32].

Nonlinear conservation laws have proven to feature many advantages in their interplay with applications (in particular in fluid dynamics and in traffic flow). First of all, they provide a very simple description of the *shock structures* in the phenomena under study (see e.g. compression waves in fluid mechanics, or queues in vehicular traffic). Most importantly, the mathematical theory developed in the past decades have greatly helped the development of efficient numerical schemes (and vice versa!), which are much handier and easier to manage when compared to agent based methods arising from a microscopic ‘moving particles’ approach. Consequently, the continuum approach based on conservation laws allows to state and possibly solve optimal control and optimal management problems, and to easily extend the theory to more complex structures (such as networks).

1.2. Goal of the paper: the many particle approximation of a scalar conservation law. The goal of this paper is to prove that (under reasonable assumptions on the velocity map v , and assuming nonnegative initial data) a scalar nonlinear conservation law of the form (1.1) can be *solved* as a *many particle limit* of a discrete (microscopic) model of interacting particle systems solving a suitable set of ODEs. We emphasise here that the velocity map $\rho \mapsto v(\rho)$ will be assumed to be *monotonically decreasing* w.r.t. $\rho \geq 0$, and with $v(0) = v_{\max} < +\infty$. The case of a monotonically increasing v can be recovered easily by simple modifications in our construction.

Our approach can be sketched as follows. We fix in $\mathbf{L}^1 \cap \mathbf{L}^\infty$ an initial density $\bar{\rho} \geq 0$ with compact support and having total (linear) mass $L > 0$. For a given integer $N > 0$, we split the subgraph of $\bar{\rho}$ in N adjacent regions of equal mass $\ell \doteq L/N$, with the endpoints of each region positioned at $\bar{x}_i \in \mathbb{R}$, $i = 0, \dots, N$. The points \bar{x}_i are interpreted as (ordered) particles with mass ℓ , and they are taken as initial condition to an ODE system describing the evolution of the particles in the discrete setting, namely to the *Follow-The-Leader (FTL) system*

$$\dot{x}_N(t) = v_{\max}, \quad (1.3a)$$

$$\dot{x}_i(t) = v\left(\frac{\ell}{x_{i+1}(t) - x_i(t)}\right), \quad i = 0, \dots, N-1. \quad (1.3b)$$

The points $x_i(t)$ are interpreted as moving particles on the real line. Basically, each particle moves with a velocity which is computed (through the map v) via the discrete density obtained by detecting the distance with the nearest right neighborhood of x_i at time t . The last particle on the right x_N has no other particles to its right, therefore the value of the discrete density on the right of x_N is zero, and this justify the ODE (1.3a) for x_N .

We remark here that no collisions occur between the particles (hence overtaking is not allowed), as the distance between two consecutive points can be proven to satisfy an efficient lower bound by a fixed multiple of ℓ depending on the initial condition. We shall focus on this issue in Lemma 2.1 below. We shall describe the FTL model in detail below in Section 2.2.

After having solved (1.3) for all times, we consider the empirical measure

$$\rho^N(t) = \ell \sum_{i=0}^{N-1} \delta_{x_i(t)}, \quad (1.4)$$

and prove in Theorem 2.3 that its limit (in a measure sense to be explained later on) as N goes to infinity is actually an \mathbf{L}^1 density ρ , which satisfies the scalar conservation law (1.1) with $f(\rho) = \rho v(\rho)$ in the Oleinik-Kružkov entropy sense [32, 40], see Definition 2.1 below.

Our convergence result has a natural interpretation as a *many particle limit* for the scalar conservation law (1.1). In this sense, it can be seen as an abstract particle method for (1.1) which can be also applied in

the context of numerics. On the other hand, the discrete model (1.3) can be also interpreted as a discrete *Lagrangian* formulation of (1.1), which makes our result meaningful from a physical point of view, as it validates the use of the macroscopic model (1.1) in cases in which the microscopic dynamics are easier to justify compared to the macroscopic ones.

Although the literature on nonlinear conservation laws is extremely rich of effective numerical schemes (we mention here the pioneering work of Glimm [25] for systems, and the wave-front tracking algorithm proposed by Dafermos in [17] and improved later on by Di Perna [22] and Bressan [10], see [11] and the references therein for more details), to our knowledge the rigorous approximation of an entropy solution to a scalar conservation law by the empirical solution to an ODE system of Lagrangian particles in the spirit of (1.3) has not been covered yet. The recent paper [15] provides preliminary results, but it does not contain the needed estimates to justify the limiting procedure. The main novelty with respect to previous results in the literature is that our result is purely *constructive*, in the sense that it can be considered as an alternative tool to actually *solve* a scalar conservation law. No property of the limiting solution is used, except the uniqueness of entropy solutions in [32, 40] which is used to prove that the scheme has a unique limit. As a byproduct of our work, the Kruzkov entropy condition that allows to single out a unique solution to (1.1) can be now also intuitively justified by the fact that it is satisfied by our discrete particle approximation, and is inherited by the density ρ obtained in the many particle limit.

Our approach differs from most of the numerical approaches to the solution to a scalar conservation law in that it interprets the microscopic limit as a *mean field limit* of a system of interacting particles with nearest neighbour type interaction, in the spirit of (locally and non-locally) interacting particles systems in statistical mechanics, probability, kinetic theory, mathematical biology, etc. In this sense, our result can be expressed in the framework of large (deterministic) particle limits with application to several contexts in fluid mechanics, see e.g. the classical references [23, 37, 41]. In one space dimension, a key result in the context of deterministic approximations is the one by Russo [48], which applies to the linear diffusion equation, in which the diffusion operator is replaced by a nearest neighbour interaction term (see also later generalizations to nonlinear diffusion in [35]). A recent result which uses the same approach to nonlinear drift diffusion equations is presented in [36]. We also mention here the paper by Brenier and Grenier [8], which provides a particle justification of the pressureless Euler system (and a particle approximation for a scalar conservation law, although with a completely different approach and interpretation). Our approach can be considered more in the spirit of [48], applied to a scalar conservation law of traffic type.

The existing numerical method for scalar conservation laws which most resembles our particle method is probably the wave-front tracking algorithm, in which the solution is approximated by a piecewise constant profile which is discontinuous on a finite number of moving fronts. Such a structure suggests the *total variation* as the natural quantity to look at in order to perform efficient uniform estimates, and the space \mathbf{L}^1 as the natural environment to set up the problem and to measure the error in the approximation procedure. In our case, the approximating sequence is a *linear combination of Dirac's deltas*. Therefore, a *measure topology* is needed to compare the approximating solution and its limit. We shall show that the most natural choice for such a topology is (a scaled version of) the 1-*Wasserstein distance*, see [1, 52].

We emphasize that the particle approach presented here, as well as the methodology used, can be potentially adapted to detect more general macroscopic models by refining the microscopic formulation of the FTL type model. Moreover, simple boundary conditions can be achieved in the limit by simple modifications of the discrete model (e.g. with entrances and exits).

1.3. Formal derivation of the scheme: the use of Wasserstein distance. The main advantage in using the topology induced by the Wasserstein distance (in our one-dimensional context) relies on its identification with the \mathbf{L}^1 -topology in the space of *pseudo-inverses of cumulative distributions*. Roughly speaking, let ρ be a nonnegative solution to (1.1) with mass $L > 0$, and let

$$F(t, x) \doteq \int_{-\infty}^x \rho(t, x) \, dx \in [0, L],$$

be its primitive. The pseudo inverse variable

$$X(t, z) \doteq \inf \{x \in \mathbb{R} : F(x) > z\}, \quad z \in [0, L[,$$

formally satisfies the *Lagrangian PDE*

$$X_t(t, z) = v \left(\frac{1}{X_z(t, z)} \right).$$

Now, if we replace the z -derivative of X by the forward finite difference

$$X_z \approx \frac{X(t, z + \ell) - X(t, z)}{\ell},$$

and assume that X is piecewise constant on intervals of length ℓ , the ODE system (1.3) is immediately recovered, with the structure

$$X(t, z) = \sum_i x_i(t) \chi_{[i\ell, (i+1)\ell[}(z).$$

We shall explain the above formal computation more in detail in Section A in the Appendix.

The use of pseudo-inverse variables and Wasserstein distances in the framework of scalar conservation laws is not totally new. In [12], a contraction estimate in the so-called ∞ -Wasserstein distance for genuinely nonlinear scalar conservation laws was derived. The case of non-decreasing solution was treated earlier in [7]. In the special case of the LWR equation for traffic flow, we also remark here that in [38] a simplified version of the model (1.1) is derived by introducing as new variable the cumulative number of vehicles passing through a location x at time t starting from the passage of some reference vehicle, see [2, 19] for recent developments of this theory.

1.4. Technical aspects of the problem. The discrete Oleinik condition. From the technical point of view, our convergence result relies first of all on proving that the empirical measure (1.4) has the same (weak) $N \rightarrow +\infty$ limit as the piecewise constant approximation

$$\hat{\rho}^N(t, x) = \sum_{i=1}^{N-1} y_i(t) \chi_{[x_i(t), x_{i+1}(t)[}, \quad y_i(t) \doteq \frac{\ell}{x_{i+1}(t) - x_i(t)},$$

in which $y_i(t)$ is the discrete Lagrangian version of the density. The most important step, however, lies in providing strong \mathbf{L}^1 compactness of $\hat{\rho}^N$. This task is performed in two different ways. In the case of **BV** initial data, we are able to provide a direct estimate of the total variation of the discrete density (see Proposition 3.5). On the other hand, our main result concerns with the case of general \mathbf{L}^∞ data: in this case, a key estimate on the particle model (see Lemma 3.4), which can be considered as a discrete version of the Oleinik condition for the scalar conservation law, allows to provide strong compactness even if the initial total variation is unbounded. In some sense, this proves that the one-sided Lipschitz regularizing effect of the scalar conservation law (1.1) is somehow an intrinsic property of the discrete Lagrangian formulation of the model. We defer to [26] and the references therein for general results on the regularizing effect for scalar conservation laws.

For numerical purposes, the use of discrete Oleinik conditions has been addressed before for the Lax-Friedrichs and Godunov schemes in [9, 27, 40, 50]. There is also a similar result for second order systems in [6]. The striking novelty in our approach is the fact that our discrete Oleinik condition is only posed in terms of the velocity field, whereas the classical Oleinik condition is stated in terms of the derivative of the flux, see [30]. This is due to the fact that the discrete model is a Lagrangian one, and is therefore characterised by the velocity law. The advantage of having the discrete one-sided Lipschitz condition in terms of the velocity is that we can also consider velocity laws with degenerate slopes at $\rho = 0$. An interesting numerical feature (which is however quite natural when considering particle approximations) is that the discrete approximation $\hat{\rho}^N$ for the density has *no vacuum regions* in the interior of its support, no matter whether or not the (continuum) initial condition is made up by more than one hump. Finally, let us mention

that our discrete density $\hat{\rho}^N$ is always discontinuous on at most $N+1$ fronts, unlike in the wave front tracking approximation in which the number of jumps may increase in time.

1.5. A motivating example: traffic flow. The macroscopic variables describing vehicular traffic are the (mean) density ρ (number of vehicles per unit length of the road), the (mean) velocity v (space covered per unit time by the vehicles) and the flux f (number of vehicles per unit time). By definition we have that

$$f = \rho v.$$

The conservation of the number of vehicles along a road with no entrance or exit is expressed by the PDE

$$\rho_t + f_x = 0,$$

where t is the time and x the position along the road. To close the above system of two equations and three unknowns, a further condition has to be imposed. Two main approaches are used in the literature (see [31, 42, 45] for a survey): first order and second order models. The former are based on a constitutive law, which expresses one of the three unknowns as function of the remaining two. In the latter, a further evolution equation is imposed. The prototype of first order models is the Lighthill-Whitham-Richards model (LWR) [34, 44], which assumes that the velocity can be expressed as an explicit function of the density alone. Hence, LWR is nothing but (1.1)-(1.2) with the choice of v depending on the specific situation, see Example 2.4 below. The most celebrated second order model is the Aw-Rascle-Zhang model (ARZ) [4, 53], which adds an evolution equation that can be regarded as a continuum analogue of Newton's law.

The continuum assumption is not immediately justifiable in the context of vehicular traffic, as the number of vehicles is typically far lower than the typical number of molecules e.g. in fluid dynamics. Usually, the continuum hypothesis is accepted as a technical approximation of the physical reality, regarding macroscopic quantities as measures of traffic features. In order to justify and make more clear the continuum hypothesis, the study of the discrete-to-continuum limit for second order models has been proposed in [3, 6]. First attempts at analyzing the particle approximation for first order macroscopic models have been recently proposed in [14, 15, 46]. From the modelling point of view, the discrete-to-continuum limit in the context of traffic flow may be also considered as the theoretical counterpart of reconstructing the traffic state of a region through high-sampling data from GPS devices. Continuum traffic models can be also detected from the microscopic scale delivered by kinetic models, see e.g. the approach proposed in [20] leading to a variant of the model of [4]. We mention also the recent [5], in which interactions at the microscopic scale are modeled by methods of game theory, thus leading to the derivation of mathematical models within the framework of the kinetic theory.

The FTL approximation scheme has been also used to approximate second order models for vehicular traffic such as ARZ [4, 53]. An important result in this sense is the one in [3]. However, differently from [3], in our result we do not shrink the length of the vehicles to zero and we do not let the size of the highway or the number of vehicles under consideration tend to infinity. In fact, our approximation algorithm rather lets the number of particles (platoons of possible fractional vehicles) under consideration tend to infinity, but keeps both the length of the domain (highway) and the total mass (total number of vehicles) L constant. Finally, another important difference from [3] is that our approach allows to handle *vacuum regions*. This introduces further technical difficulties that are rigourously treated and solved in the present paper.

1.6. Structure of the paper. The present paper is structured as follows. Section 2 contains all the preliminary material and the statement of the main result. More precisely, we summarise the mathematical theory of the one-dimensional scalar conservation law (1.1) in Section 2.1, we define the (discrete) FTL model (1.3) in detail in Section 2.2. we recall the basics on the Wasserstein distance in one space dimension in Section 2.3, and we set up the approximating scheme and state our main result in Section 2.4. The precise statement of the main result is contained in Theorem 2.3. Section 3 is devoted to its proof, and is split into the subsections 3.1, 3.2, 3.3, and 3.4. More precisely, Section 3.1 is devoted to the proof of the weak convergence of our approximating scheme, in Section 3.2 we prove the two basic compactness estimates mentioned above, in Section 3.3 we provide the needed time-continuity and prove strong compactness in \mathbf{L}^1 , and finally in Section 3.4

we prove that the limit of our approximating scheme is the unique entropy solution in the Oleinik-Kružkov entropy sense [32, 40].

2. PRELIMINARIES AND STATEMENT OF THE MAIN RESULT

2.1. One-dimensional scalar conservation law. Consider the Cauchy problem for a scalar conservation law in one space dimension

$$\rho_t + [\rho v(\rho)]_x = 0, \quad t > 0, \quad x \in \mathbb{R}, \quad (2.1a)$$

$$\rho(0, x) = \bar{\rho}(x), \quad x \in \mathbb{R}, \quad (2.1b)$$

where $\rho = \rho(t, x) \geq 0$ represents the density of the medium at the position $x \in \mathbb{R}$ at time $t \geq 0$, $v = v(\rho)$ is the local velocity and $\bar{\rho}$ is the initial datum, see e.g. [11, 18, 49] as general references. Regularity (and convexity) assumptions are typically required in the literature on the *flux* function $f(\rho) \doteq \rho v(\rho)$. However, *our assumptions will be on the velocity v rather than on the flux f* , see Section 2.4.

Let us briefly recall the main properties of the solutions to (2.1). If the initial datum has compact support, then the support of any solution has finite speed of propagation. The maximum principle holds true, and if the initial datum is nonnegative, then the solutions remain nonnegative for all times. Moreover, the total (linear) mass $\int_{\mathbb{R}} \rho(t, x) dx$ is time independent: $\int_{\mathbb{R}} \rho(t, x) dx = L \doteq \int_{\mathbb{R}} \bar{\rho}(x) dx$ for all $t \geq 0$.

It is well known that the solutions to (2.1) may develop discontinuities in a finite time, also for regular initial data. For this reason, one has to consider weak solutions ρ to (2.1), more precisely ρ in $\mathbf{L}^\infty([0, +\infty[; \mathbf{L}^1 \cap \mathbf{L}^\infty(\mathbb{R}))$ that satisfy (2.1) in the sense of distributions, namely

$$\int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[\rho(t, x) \varphi_t(t, x) + f(\rho(t, x)) \varphi_x(t, x) \right] dt dx + \int_{\mathbb{R}} \bar{\rho}(x) \varphi(0, x) dx = 0 \quad (2.2)$$

for all $\varphi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$. The choice of $\mathbf{L}^1 \cap \mathbf{L}^\infty(\mathbb{R})$ as the functional space to deal with the x -regularity appears as the most reasonable one in order to obtain existence of weak solutions when the approximating procedure is performed via a vanishing viscosity argument, see e.g. [18, Section 6.3]. However, the space $\mathbf{BV}(\mathbb{R})$ is more reminiscent of the typical structure of solutions featuring shocks and rarefaction waves, and turns out to be a natural choice when the problem is e.g. solved by the polygonal approximation algorithm also known as the wave-front tracking algorithm [17], see also [11] and the references therein.

It is well known that the notion of weak solution introduced above is not strong enough to provide uniqueness of solutions to (2.1). The concept of entropy solution formulated in [32, 33, 40] (see also [18] and the references therein), provides the most natural and efficient way to single out a unique (physically relevant) solution to (2.1). Such concept can be formulated in several ways, also depending on the regularity of ρ , the most general one being the one proposed by Kružkov [32], which holds for a reasonably wide class of fluxes (namely $\rho \mapsto f(\rho)$ being locally Lipschitz) and in arbitrary space dimension.

Definition 2.1 (Entropy solutions). *Assume that the flux $\rho \mapsto f(\rho)$ is locally Lipschitz. A function ρ in $\mathbf{L}^\infty([0, +\infty[; \mathbf{L}^1 \cap \mathbf{L}^\infty(\mathbb{R}))$ is an entropy solution to (2.1) if it satisfies the entropy inequality*

$$\begin{aligned} & \int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[|\rho(t, x) - k| \varphi_t(t, x) + \operatorname{sgn}(\rho(t, x) - k) [f(\rho(t, x)) - f(k)] \varphi_x(t, x) \right] dt dx \\ & + \int_{\mathbb{R}} |\bar{\rho}(x) - k| \varphi(0, x) dx \geq 0 \end{aligned} \quad (2.3)$$

for all $\varphi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$ with $\varphi \geq 0$, and for all constants $k \in \mathbb{R}$.

Clearly, any entropy solution is a weak solution to (2.1) in the sense of (2.2). Moreover uniqueness follows from (2.3).

Theorem 2.1 (Kružkov [32]). *Assume that the flux f is locally Lipschitz. Then, for any given initial condition $\bar{\rho}$ in \mathbf{L}^∞ with compact support, there exists a unique entropy solution to (2.1) in the sense of Definition 2.1.*

It is easy to check that any function ρ satisfying the entropy inequality (2.3) satisfies also the following property of (weak) \mathbf{L}^1 -continuity in time

$$\lim_{T \rightarrow 0+} \frac{1}{T} \int_0^T \int_{|x| \leq r} |\rho(t, x) - \bar{\rho}(x)| dx dt = 0$$

for all $r > 0$. However, depending on the way we attempt at constructing entropy solutions, an important issue is related with detecting the trace at $t = 0$ in a strong enough topology. This is often the case when the approximating scheme lacks of compactness when t approaches zero. A theorem due to Chen and Rascle [13] states that the uniqueness of the entropy solution is preserved also for a notion of entropy solution relaxed at $t = 0$, provided the flux f satisfies a.e. a genuine nonlinearity condition.

Theorem 2.2 (Chen and Rascle [13]). *Assume there exists no nontrivial interval on which f is affine. If $\bar{\rho}$ is in \mathbf{L}^∞ and has compact support, then there exists a unique ρ in $\mathbf{L}^\infty([0, +\infty[; \mathbf{L}^1 \cap \mathbf{L}^\infty(\mathbb{R}))$ weak solution to (2.1) in the sense of (2.2) that satisfies also*

$$\int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[|\rho(t, x) - k| \varphi_t(t, x) + \operatorname{sgn}(\rho(t, x) - k) [f(\rho(t, x)) - f(k)] \varphi_x(t, x) \right] dt dx \geq 0 \quad (2.4)$$

for all $\varphi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$ with $\varphi \geq 0$, and for all constants $k \in \mathbb{R}$. Moreover, ρ is the unique entropy solution in the sense of Definition 2.1.

Let us finally recall that, for \mathbf{C}^1 -fluxes f which are concave or convex, another classical tool to uniquely determine all weak solutions by their \mathbf{L}^∞ -initial values is the so called Oleinik-type condition [30]

$$\int_{\mathbb{R}} \int_{\mathbb{R}_+} f'(\rho(t, x)) \varphi_x(t, x) dt dx \geq - \int_{\mathbb{R}} \int_{\mathbb{R}_+} \frac{1}{t} \varphi(t, x) dt dx \quad (2.5)$$

for all $\varphi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$ with $\varphi \geq 0$, and for all $t > 0$. Moreover, if f' has Lipschitz continuous inverse, then (2.5) implies that $\rho(t, \cdot)$ has locally bounded total variation for all $t > 0$ even if the initial datum is not locally in \mathbf{BV} .

2.2. Follow-the-leader model. Microscopic models of vehicular traffic are typically based on the so called Follow-The-Leader (FTL) model. On the other hand, such model makes sense in any context in which a set of discrete ordered agents moves with a velocity computed instantaneously as a function of the discrete density.

Consider $N + 1$ ordered particles localised on \mathbb{R} . Denote by $t \mapsto x_i(t)$ the position of the i -th particle for $i = 0, \dots, N$. Then, according to the FTL model, the evolution of the particles (which mimics the evolution of the position of $N + 1$ vehicles along the road) is described inductively by the following Cauchy problem for an ODE system

$$\dot{x}_N(t) = v_{\max}, \quad (2.6a)$$

$$\dot{x}_i(t) = v \left(\frac{\ell}{x_{i+1}(t) - x_i(t)} \right), \quad i = 0, \dots, N - 1, \quad (2.6b)$$

$$x_i(0) = \bar{x}_i, \quad i = 0, \dots, N, \quad (2.6c)$$

where v is a \mathbf{C}^1 strictly decreasing velocity map, $\bar{x}_0 < \dots < \bar{x}_N$ are the initial positions of the particles, and $\ell > 0$ is the (linear) mass of each particle. We shall assume that v is strictly decreasing and bounded from above. Then, $v_{\max} \doteq v(0)$ is the maximum possible velocity and is reached only by the leading particle x_N . Let us underline that in vehicular traffic the FTL model requires that all vehicles move towards a unique direction (i.e. the vehicles move along a one way road), e.g. v is nonnegative. Our setting is more general and such an assumption is not required. Notice in particular that we shall not prescribe any constrain on the sign of v_{\max} .

We now set

$$R \doteq \max_{i=0, \dots, N-1} \left(\frac{\ell}{\bar{x}_{i+1} - \bar{x}_i} \right).$$

The quantity $R > 0$ is the *maximum discrete density* at time $t = 0$ and

$$\bar{x}_{i+1} - \bar{x}_i \geq \frac{\ell}{R}, \quad i = 0, \dots, N-1. \quad (2.7)$$

System (2.6) can be solved inductively starting from $i = N$. Indeed, from (2.6a), we immediately deduce that

$$x_N(t) = \bar{x}_N + v_{\max} t.$$

Then, we can compute $t \mapsto x_i(t)$ once we know $t \mapsto x_{i+1}(t)$. In fact, according with the system (2.6) the velocity of the i -th particle depends on its distance from the $(i+1)$ -th particle alone via the smooth velocity map v . In order to ensure that the (unique) solution to (2.6) exists globally in $t \geq 0$, we need to prove that the distances $x_{i+1}(t) - x_i(t)$ never degenerate. We can actually prove that the discrete density never exceeds the upper bound $R > 0$ holding at $t = 0$. This is proven in the next lemma, which improves a similar one in [46], which only holds in the case $v(\rho_{\max}) = 0$ for some maximal density $\rho_{\max} > 0$, and in which the authors prove that the discrete density stays bounded by the threshold density ρ_{\max} . Our result is actually much stronger, because it basically proves that the discrete model satisfies the same maximum principle of the corresponding continuum model (see e.g. [18, Theorem 6.2.4]), i.e. the discrete density at an arbitrary time $t > 0$ is controlled by the supremum norm of the initial discrete density.

Lemma 2.1 (Discrete maximum principle). *For all $i = 0, \dots, N-1$, we have*

$$\frac{\ell}{R} \leq x_{i+1}(t) - x_i(t) \leq \bar{x}_N - \bar{x}_0 + (v_{\max} - v(R))t \quad \text{for all times } t \geq 0. \quad (2.8)$$

Proof. The upper bound is obvious. Hence, it suffices to prove the lower bound. At time $t = 0$ the lower bound is satisfied because of (2.7). We shall prove that

$$\sup_{t \geq 0} [x_{j+1}(t) - x_j(t)] \geq \frac{\ell}{R}, \quad j = 0, \dots, N-1, \quad (2.9)$$

by a recursive argument on j . The statement is true for $j = N-1$. Indeed,

$$\begin{aligned} x_N(t) - x_{N-1}(t) &= \bar{x}_N - \bar{x}_{N-1} + \int_0^t \left[v_{\max} - v \left(\frac{\ell}{x_N(s) - x_{N-1}(s)} \right) \right] ds \\ &\geq \bar{x}_N - \bar{x}_{N-1} \geq \frac{\ell}{R}, \end{aligned}$$

because $v(\rho) \leq v_{\max}$ for all $\rho \geq 0$. Assume now that

$$\sup_{t \geq 0} [x_{j+2}(t) - x_{j+1}(t)] \geq \frac{\ell}{R}. \quad (2.10)$$

Assume by contradiction that there exists $j \in \{0, \dots, N-2\}$ and $t_2 > t_1 \geq 0$ such that

$$x_{j+1}(t_1) - x_j(t_1) = \frac{\ell}{R}$$

and

$$x_{j+1}(t) - x_j(t) < \frac{\ell}{R} \quad \text{for all } t \in]t_1, t_2]. \quad (2.11)$$

Since v is strictly decreasing, we have for any $t \in]t_1, t_2]$

$$x_j(t) = x_j(t_1) + \int_{t_1}^t v \left(\frac{\ell}{x_{j+1}(t) - x_j(t)} \right) dt \leq x_j(t_1) + v(R)(t - t_1). \quad (2.12)$$

By (2.10) we have for any $t \in]t_1, t_2]$

$$x_{j+1}(t) = x_{j+1}(t_1) + \int_{t_1}^t v \left(\frac{\ell}{x_{j+2}(t) - x_{j+1}(t)} \right) dt \geq x_{j+1}(t_1) + v(R)(t - t_1),$$

and therefore by (2.12)

$$x_{j+1}(t) - x_j(t) \geq x_{j+1}(t_1) - x_j(t_1) = \frac{\ell}{R},$$

which contradicts (2.11). Hence, (2.9) is satisfied and the assertion is proven. \diamond

We emphasise that the above discrete maximum principle is a direct consequence of the *transport* nature behind the FTL system (2.6). Indeed, the global bound for the discrete density is propagated from the last particle x_N back to all the other particles, as emphasised by the recursive argument in the proof of Lemma 2.1.

2.3. Notation and preliminaries on measure distances. In this section we recall basic properties of pseudo-inverse operators and on the one-dimensional Wasserstein distance that we shall use extensively in the rest of the paper. We defer to [52] for further details.

For a fixed $L > 0$, introduce the pseudo-inverse operators

$$\begin{aligned} \mathcal{X} : \mathbf{L}^\infty(\mathbb{R}; [0, L]) &\rightarrow \mathbf{L}^\infty([0, L[; \mathbb{R}), \\ \mathcal{F} : \mathbf{L}^\infty([0, L[; \mathbb{R}) &\rightarrow \mathbf{L}^\infty(\mathbb{R}; [0, L]), \end{aligned}$$

defined by

$$\begin{aligned} \mathcal{X}[F](z) &\doteq \inf \{x \in \mathbb{R} : F(x) > z\} && \text{for } z \in [0, L[, \\ \mathcal{F}[X](x) &\doteq \text{meas} \{z \in [0, L] : X(z) \leq x\} && \text{for } x \in \mathbb{R}, \end{aligned}$$

and consider the space

$$\mathcal{M}_L \doteq \{\rho \text{ Radon measure on } \mathbb{R} \text{ with compact support} : \rho \geq 0, \rho(\mathbb{R}) = L\}. \quad (2.13)$$

For a given $\rho \in \mathcal{M}_L$, we denote $x_{\min}^\rho \doteq \min(\text{spt}(\rho))$ and $x_{\max}^\rho \doteq \max(\text{spt}(\rho))$, and by $F_\rho : \mathbb{R} \rightarrow [0, L]$ its cumulative distribution, namely

$$F_\rho(x) \doteq \rho([-\infty, x]).$$

We observe that $F_\rho \in \mathbf{L}^\infty(\mathbb{R}; [0, L])$ is non-decreasing, right-continuous with $F_\rho(x) = 0$ for all $x < x_{\min}^\rho$ and $F_\rho(x) = L$ for all $x \geq x_{\max}^\rho$. Therefore we can define its pseudo-inverse $X_\rho \doteq \mathcal{X}[F_\rho]$. Clearly, $X_\rho \in \mathbf{L}^\infty([0, L[; [x_{\min}^\rho, x_{\max}^\rho])$ is non-decreasing, right-continuous with $X_\rho(0) = x_{\min}^\rho$. By abuse of notation, we shall adopt the notation ρ to denote an absolutely continuous measure in \mathcal{M}_L with \mathbf{L}^1 -density ρ . We recall the following lemma (see e.g. [52]).

Lemma 2.2 (Change of variable). *If $\rho \in \mathcal{M}_L$, then for all $\varphi \in \mathbf{C}^0(\mathbb{R}; \mathbb{R})$ we have*

$$\int_{\mathbb{R}} \varphi(x) d\rho(x) = \int_0^L \varphi(X_\rho(z)) dz.$$

We recall that, for $L = 1$, the one-dimensional 1-Wasserstein distance between $\rho_1, \rho_2 \in \mathcal{M}_1$ (defined in terms of optimal plans in the Monge-Kantorovich problem, see e.g. [52]) can be defined as

$$d_1(\rho_1, \rho_2) \doteq \|F_{\rho_1} - F_{\rho_2}\|_{\mathbf{L}^1(\mathbb{R}; \mathbb{R})} = \|X_{\rho_1} - X_{\rho_2}\|_{\mathbf{L}^1([0, 1]; \mathbb{R})}.$$

For a general strictly positive L , we introduce the *scaled 1-Wasserstein distance* between $\rho_1, \rho_2 \in \mathcal{M}_L$ as

$$d_{L,1}(\rho_1, \rho_2) \doteq \|F_{\rho_1} - F_{\rho_2}\|_{\mathbf{L}^1(\mathbb{R}; \mathbb{R})} = \|X_{\rho_1} - X_{\rho_2}\|_{\mathbf{L}^1([0, L]; \mathbb{R})}. \quad (2.14)$$

Indeed, straightforward computation yields

$$d_{L,1}(\rho_1, \rho_2) = L d_1(\rho_1/L, \rho_2/L).$$

The distance $d_{L,1}$ inherits all the topological properties of the 1-Wasserstein distance for probability measures. In particular, a sequence $(\rho_n)_{n \in \mathbb{N}}$ in \mathcal{M}_L converges to $\rho \in \mathcal{M}_L$ in $d_{L,1}$ if and only if

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \varphi(x) d\rho_n(x) = \int_{\mathbb{R}} \varphi(x) d\rho(x),$$

for all $\varphi \in \mathbf{C}^0(\mathbb{R}; \mathbb{R})$ growing at most linearly at infinity.

2.4. Statement of the main result. In this subsection we state our main result, which provides a rigorous description of the unique entropy solution ρ to the Cauchy problem (2.1) as the limit for N that goes to infinity of a density associated to the microscopic model (2.6) to be constructed as described below.

We shall work under the standing assumption on the initial datum

(In) The initial datum $\bar{\rho}$ is in $\mathcal{M}_L \cap \mathbf{L}^\infty(\mathbb{R})$,

where \mathcal{M}_L is defined in (2.13). In some cases we shall require the stronger condition

(InBV) The initial datum $\bar{\rho}$ is in $\mathcal{M}_L \cap \mathbf{BV}(\mathbb{R})$.

As for the velocity function v , we shall require throughout the paper:

(V1) $v \in \mathbf{C}^1([0, +\infty[)$, v strictly decreasing on $[0, +\infty[$.

(V2) $v(0) = v_{\max}$ for some $v_{\max} \in \mathbb{R}$.

The \mathbf{C}^1 assumption in (V1) is a minimal requirement for having a unique local solution to the system (2.6). The monotonicity assumption in (V1) is a natural one in traffic models. However, we shall see that it is technically relevant in many parts of our paper, in particular in Section 3.2. The assumption (V2) basically states that $v(0)$ is finite. We notice that the maximum principle for (2.1) and the discrete maximum principle in Lemma 2.1 imply, together with the monotonicity of v , that both $v(\rho)$ and $v(\ell/(x_{i+1}(t) - x_i(t)))$ are globally bounded in time by a constant depending on the \mathbf{L}^∞ norm of the initial datum. The assumption $v(0) < +\infty$ is a natural requirement in order to guarantee the existence of a solution to the FTL system (2.6). Indeed, if $v(0) = +\infty$, then the particle x_N goes to $+\infty$ instantaneously after $t = 0$, and so do all other particles. Therefore, the particle system cannot approximate the continuum equation (2.6) in the many particle limit, and the assumption $v(0) < +\infty$ is a necessary condition for our result. Let us also remark that the assumption (V1) implies that $f(\rho) = \rho v(\rho)$ is not affine, and hence satisfies the hypotheses of Theorem 2.2.

In some cases, we shall use the stronger assumption

(V3) The map $[0, +\infty[\ni \rho \mapsto \rho v'(\rho) \in [0, +\infty[$ is non increasing.

Notice that the assumption (V3) implies in particular that the flux $\rho \mapsto f(\rho) = \rho v(\rho)$ is concave. Indeed, since $f'(\rho) = v(\rho) + \rho v'(\rho)$, f' is the sum of two non-increasing functions if (V3) is satisfied. The assumption (V3) is clearly a stricter requirement than the concavity of f ; on the other hand, (V3) is verified by many examples of velocities arising in traffic flow models.

Example. [Examples of velocities in vehicular traffic] In vehicular traffic, a maximal density $\rho_{\max} > 0$ is prescribed, at which all vehicles are bumper to bumper. Typically, ρ in (2.1) represents a normalised density, and one can assume $\rho_{\max} = 1$. The prototype for the velocity in vehicular traffic $v(\rho) = v_{\max}(1 - \rho)$ by Greenshields [29] clearly satisfies the assumptions (V1), (V2), (V3). The same holds for the Pipes-Munjial velocity [43]

$$v(\rho) = v_{\max}(1 - \rho^\alpha) \quad \alpha > 0,$$

in which the concavity of the flux $\rho v(\rho)$ degenerates at $\rho = 0$, and for the Underwood model [51]

$$v(\rho) = v_{\max} e^{-\rho}.$$

A further example of speed-density relations that satisfy (V1), (V2), (V3) are

$$v(\rho) = v_{\max} \left[\log \left(\frac{1}{\alpha} \right) \right]^{-1} \log \left(\frac{1}{\rho + \alpha} \right), \quad \alpha > 0,$$

that result from a slight modification of the Greenberg model [28]

We now introduce our atomization scheme. Denote by $\bar{x}_{\min} < \bar{x}_{\max}$ the extremal points of the convex hull of the support of $\bar{\rho}$, namely $\bigcap_{[a,b] \supseteq \text{spt}(\bar{\rho})} [a,b] = [\bar{x}_{\min}, \bar{x}_{\max}]$. Fix $n \in \mathbb{N}$ sufficiently large. Let $L \doteq \|\bar{\rho}\|_{\mathbf{L}^1(\mathbb{R})} > 0$ and split the subgraph of $\bar{\rho}$ in $N_n \doteq 2^n$ regions of measure $\ell_n \doteq 2^{-n}L$ as follows. Set

$$\bar{x}_0^n \doteq \bar{x}_{\min}, \quad (2.15a)$$

and recursively

$$\bar{x}_i^n \doteq \sup \left\{ x \in \mathbb{R} : \int_{\bar{x}_{i-1}^n}^x \bar{\rho}(y) \, dy < \ell_n \right\}, \quad i = 1, \dots, N_n. \quad (2.15b)$$

It is easily seen that $\bar{x}_{N_n}^n = \bar{x}_{\max}$, and $\bar{x}_{N_n-i}^n = \bar{x}_{N_n+m-2^m i}^{n+m}$ for all $i = 0, \dots, N_n$. Since we are always assuming that $\bar{\rho} \in \mathbf{L}^\infty(\mathbb{R})$, let us set

$$R \doteq \|\bar{\rho}\|_{\mathbf{L}^\infty(\mathbb{R})}.$$

We have

$$\ell_n = \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \bar{\rho}(y) \, dy \leq (\bar{x}_{i+1}^n - \bar{x}_i^n) R, \quad i = 0, \dots, N_n - 1.$$

Thus the condition (2.7) is satisfied with $\ell = \ell_n$, and we can take the values $\bar{x}_0^n, \dots, \bar{x}_{N_n}^n$ as the initial positions of the $(N_n + 1)$ particles in the n -depending version of the FTL model (2.6)

$$\dot{x}_{N_n}^n(t) = v_{\max}, \quad (2.16a)$$

$$\dot{x}_i^n(t) = v \left(\frac{\ell_n}{x_{i+1}^n(t) - x_i^n(t)} \right), \quad i = 0, \dots, N_n - 1, \quad (2.16b)$$

$$x_i^n(0) = \bar{x}_i^n, \quad i = 0, \dots, N_n. \quad (2.16c)$$

The existence of a global-in-time solution to (2.16) follows from Lemma 2.1. Moreover, from (2.16a) we immediately deduce that

$$x_{N_n}^n(t) = \bar{x}_{\max} + v_{\max} t.$$

Finally, since v is decreasing, and its argument $\ell/[x_{i+1}^n(t) - x_i^n(t)]$ is always bounded above by R , we have

$$x_0^n(t) \geq \bar{x}_{\min} + v(R) t.$$

We stress once again that $v(R)$ may be negative.

By introducing in (2.16) the new variable

$$y_i^n(t) \doteq \frac{\ell_n}{x_{i+1}^n(t) - x_i^n(t)}, \quad i = 0, \dots, N_n - 1, \quad (2.17)$$

we obtain

$$\dot{y}_{N-1}^n = -\frac{(y_{N-1}^n)^2}{\ell_n} [v_{\max} - v(y_{N-1}^n)], \quad (2.18a)$$

$$\dot{y}_i^n = -\frac{(y_i^n)^2}{\ell_n} [v(y_{i+1}^n) - v(y_i^n)], \quad i = 0, \dots, N_n - 2, \quad (2.18b)$$

$$y_i^n(0) = \bar{y}_i^n \doteq \frac{\ell_n}{\bar{x}_{i+1}^n - \bar{x}_i^n}, \quad i = 0, \dots, N_n - 1. \quad (2.18c)$$

Observe that $\ell_n/[\bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R))t] \leq y_i^n(t) \leq R$ for all $t \geq 0$ in view of Lemma 2.1. The quantity y_i^n can be seen as a discrete version of the density ρ in Lagrangian coordinates, and the ODEs (2.18a)–(2.18b) are a discrete Lagrangian version of the scalar conservation law (2.1a).

We are now ready to state the main result of this paper.

Theorem 2.3. *Let $\bar{\rho}$ satisfy the condition (In) and v the condition (V1) and (V2). Assume further that either*

- $\bar{\rho}$ satisfies (InBV),

or

- v satisfies (V3).

Define the piecewise constant (with respect to x) density

$$\hat{\rho}^n(t, x) \doteq \sum_{i=0}^{N_n-1} y_i^n(t) \chi_{[x_i^n(t), x_{i+1}^n(t)]}(x), \quad (2.19)$$

and the empirical measure

$$\tilde{\rho}^n(t, x) \doteq \ell_n \sum_{i=0}^{N_n-1} \delta_{x_i^n(t)}(x). \quad (2.20)$$

Then the sequence $(\hat{\rho}^n)_{n \in \mathbb{N}}$ converges to the unique entropy solution ρ of the Cauchy problem (2.1) almost everywhere and in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times \mathbb{R})$. Moreover, the sequence $(\tilde{\rho}^n)_{n \in \mathbb{N}}$ converges to ρ in the topology of $\mathbf{L}_{\text{loc}}^1([0, +\infty[; d_{L,1})$.

3. PROOF OF THE MAIN RESULT

Our strategy for the proof of Theorem 2.3 can be resumed as follows:

- (i) Following the notation introduced in Section 2.3, we set $\hat{F}^n = F_{\hat{\rho}^n}$ and $\hat{X}^n = \mathcal{X}[\hat{F}^n]$, respectively $\tilde{F}^n = F_{\tilde{\rho}^n}$ and $\tilde{X}^n = \mathcal{X}[\tilde{F}^n]$, as the cumulative distribution of $\hat{\rho}^n$, respectively $\tilde{\rho}^n$, and its pseudo inverses. Introduce the *discrete Lagrangian density*

$$\check{\rho}^n = \hat{\rho}^n \circ \hat{X}^n.$$

- (ii) As a first step we prove that the sequence of piecewise constant pseudo-inverse distributions $(\tilde{X}^n)_{n \in \mathbb{N}}$ has a strong limit X in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L]; \mathbb{R})$, which is equivalent to having $(\tilde{\rho}^n)_{n \in \mathbb{N}}$ converging to a measure ρ in the $\mathbf{L}_{\text{loc}}^1([0, +\infty[; d_{L,1})$ topology. At the same time, we shall also prove that $(\hat{X}^n)_{n \in \mathbb{N}}$ converges strongly in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L]; \mathbb{R})$ to the same limit X , i.e. $(\hat{\rho}^n)_{n \in \mathbb{N}}$ converges to ρ in $\mathbf{L}_{\text{loc}}^1([0, +\infty[; d_{L,1})$.
- (iii) We then prove that the limit pseudo-inverse function X has difference quotients bounded below by $1/R$. This fact allows to prove that the limit measure ρ in (ii) is actually in \mathbf{L}^∞ and is a.e. bounded by R . At the same time, we easily infer weak- $*$ convergence of $(\check{\rho}^n)_{n \in \mathbb{N}}$ to a limit $\check{\rho}$ in \mathbf{L}^∞ . It remains to prove that $\check{\rho} \circ F = \rho$, and that such limit is the unique entropy solution to (2.1). This requires stronger estimates on $\hat{\rho}^n$.
- (iv) A direct proof of a uniform **BV** estimate for $\hat{\rho}^n$ can be performed in the case of $\mathcal{M}_L \cap \mathbf{BV}$ initial datum. In the case of general $\mathcal{M}_L \cap \mathbf{L}^\infty$ initial datum, and with v satisfying (V3), we shall prove that the discrete Lagrangian density $\check{\rho}^n$ satisfies a (uniform) *discrete version of the Oleinik condition*, which implies automatically a **BV** uniform estimate for $\check{\rho}^n$, and hence for $\hat{\rho}^n$.
- (v) The definition of weak solution (2.2) for ρ follows from the $n \rightarrow +\infty$ limit of the formulation of (2.16) as a PDE

$$\tilde{X}_t^n = v(\check{\rho}^n). \quad (3.1)$$

- (vi) We finally recover the entropy condition (2.3) in the discrete setting, and use the strong \mathbf{L}^1 compactness to pass it to the limit.

3.1. Weak convergence of the approximating scheme. Throughout this subsection we shall assume that v satisfies (V1) and (V2). Let $\hat{\rho}^n$ and $\tilde{\rho}^n$ be defined as in (2.19) and (2.20) respectively. We have that $\hat{\rho}^n(t), \tilde{\rho}^n(t) \in \mathcal{M}_L$ for all $t \geq 0$. Thus we can consider the cumulative distributions associated to $\hat{\rho}^n$ and $\tilde{\rho}^n$ (recall that $\tilde{\rho}^n$ is an empirical measure)

$$\hat{F}^n(t, x) \doteq \int_{-\infty}^x \hat{\rho}^n(t, y) dy, \quad \tilde{F}^n(t, x) \doteq \tilde{\rho}^n([-\infty, x]),$$

and their pseudo-inverses

$$\hat{X}^n \doteq \mathcal{X}[\hat{F}^n], \quad \tilde{X}^n \doteq \mathcal{X}[\tilde{F}^n],$$

extended to $z = L$ by taking $\hat{X}^n(t, L) = x_{N_n}^n(t) = \tilde{X}^n(t, L)$. By definition, see figures 1 and 2, for all $t \geq 0$,

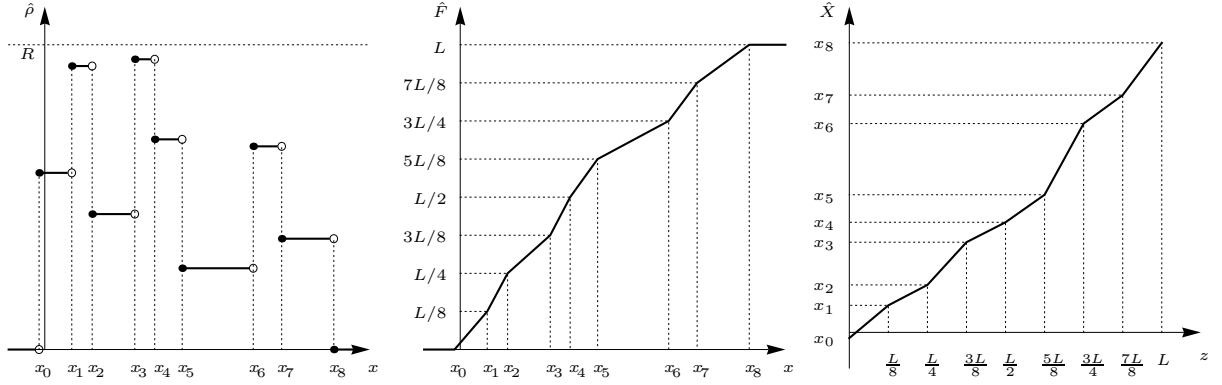


FIGURE 1. Maps of the form, respectively from the left, (2.19), (3.2) and (3.3) with $n(= 3)$ and $t(\geq 0)$ omitted.

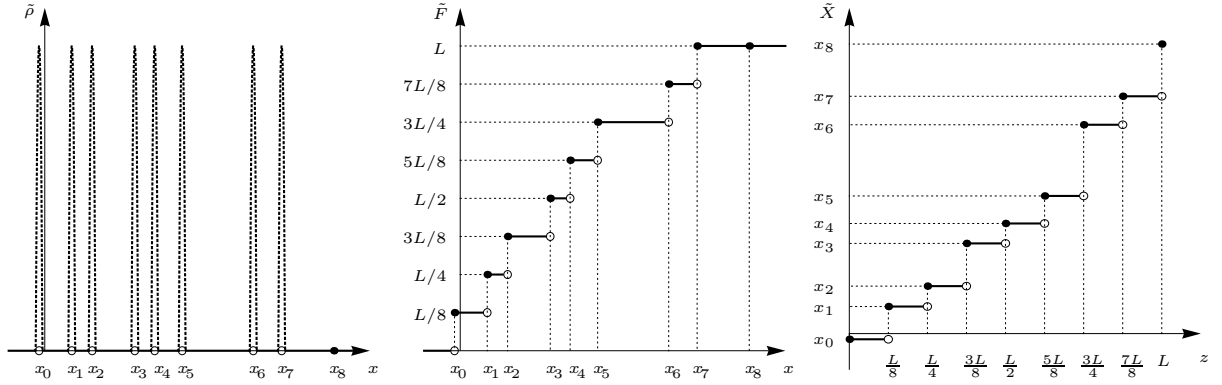


FIGURE 2. Maps of the form, respectively from the left, (2.20), (3.4) and (3.5) with $n(= 3)$ and $t(\geq 0)$ omitted.

$z \in [0, L]$ and $x \in \mathbb{R}$ we have

$$\hat{F}^n(t, x) = \sum_{i=0}^{N_n-1} \left[i \ell_n + y_i^n(t) [x - x_i^n(t)] \right] \chi_{[x_i^n(t), x_{i+1}^n(t)]}^{(x)}$$

$$+ L \chi_{[x_{N_n}^n(t), +\infty[}(x), \quad (3.2)$$

$$\begin{aligned} \hat{X}^n(t, z) &= \sum_{i=0}^{N_n-2} \left[x_i^n(t) + \frac{z - i \ell_n}{y_i^n(t)} \right] \chi_{[i \ell_n, (i+1) \ell_n[}(z) \\ &+ \left[x_{N_n-1}^n(t) + \frac{z - L + \ell_n}{y_{N_n-1}^n(t)} \right] \chi_{[L - \ell_n, L]}(z), \end{aligned} \quad (3.3)$$

$$\tilde{F}^n(t, x) = \sum_{i=0}^{N_n-2} \ell_n (i+1) \chi_{[x_i^n(t), x_{i+1}^n(t)]}(x) + L \chi_{[x_{N_n-1}^n(t), +\infty[}(x), \quad (3.4)$$

$$\tilde{X}^n(t, z) = \sum_{i=0}^{N_n-1} x_i^n(t) \chi_{[i \ell_n, (i+1) \ell_n[}(z) + x_{N_n}^n(t) \chi_{\{L\}}(z). \quad (3.5)$$

Observe that for any fixed $t \geq 0$

- $z \mapsto \hat{X}^n(t, z)$ and $x \mapsto \hat{F}^n(t, x)$ are piecewise linear continuous and non-decreasing;
- $\hat{F}^n(t) : [x_0^n(t), x_{N_n}^n(t)] \rightarrow [0, L]$ and $\hat{X}^n(t) : [0, L] \rightarrow [x_0^n(t), x_{N_n}^n(t)]$ are strictly increasing and are inverse functions of each other in the classical sense;
- $z \mapsto \tilde{X}^n(t, z)$ and $x \mapsto \tilde{F}^n(t, x)$ are piecewise constant with N_n jumps of discontinuity, right continuous and non-decreasing;
- $\hat{F}^n(t, x) \leq \tilde{F}^n(t, x)$ for any $x \in \mathbb{R}$ and $\tilde{X}^n(t, z) \leq \hat{X}^n(t, z)$ for any $z \in [0, L]$;
- $\tilde{F}^{n+1}(t, x) \leq \tilde{F}^n(t, x)$ for any $x \in \mathbb{R}$ and $\tilde{X}^n(t, z) \leq \tilde{X}^{n+1}(t, z)$ for any $z \in [0, L]$;
- $\hat{\rho}^n(t, x) = \hat{F}_x^n(t, x)$ for all $x \neq x_i^n(t)$, $i = 1, \dots, N_n$, while $\hat{\rho}^n = \tilde{F}_x^n$ in the sense of distributions.

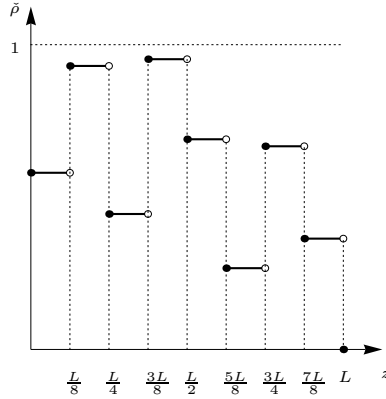


FIGURE 3. Map of the form (3.6) with $n(=3)$ and $t(\geq 0)$ omitted.

For later use, see Figure 3, we introduce also the *discrete Lagrangian density*

$$\check{\rho}^n(t, z) \doteq \hat{\rho}^n \left(t, \hat{X}^n(t, z) \right) = \sum_{i=0}^{N_n-1} y_i^n(t) \chi_{[i \ell_n, (i+1) \ell_n[}(z) \quad (3.6)$$

and observe that

$$\tilde{X}_t^n(t, z) = v(\check{\rho}^n(t, z)), \quad t > 0, \quad z \in [0, L]. \quad (3.7)$$

As a first step, we want to prove that $(\tilde{X}^n)_{n \in \mathbb{N}}$ and $(\hat{X}^n)_{n \in \mathbb{N}}$ have the same unique limit in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L]; \mathbb{R})$.

Proposition 3.1 (Definition of X). *There exists a unique non-decreasing and z -right continuous function $X = X(t, z) \in \mathbf{L}^\infty([0, +\infty[\times [0, L]; \mathbb{R})$, such that*

$$(\hat{X}^n)_{n \in \mathbb{N}} \text{ and } (\tilde{X}^n)_{n \in \mathbb{N}} \text{ converge to } X \text{ in } \mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L]; \mathbb{R}),$$

and for any $t, s > 0$

$$\text{TV}[X(t)] \leq \bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R))t, \quad (3.8a)$$

$$\|X(t)\|_{\mathbf{L}^\infty([0, L]; \mathbb{R})} \leq \max\{|\bar{x}_{\min} + v(R)t|, |\bar{x}_{\max} + v_{\max}t|\}, \quad (3.8b)$$

$$\int_0^L |X(t, z) - X(s, z)| dz \leq \max\{|v_{\max}|, |v(R)|\} L |t - s|. \quad (3.8c)$$

Moreover, $(\tilde{X}^n)_{n \in \mathbb{N}}$ converges to X a.e. on $[0, +\infty[\times [0, L]$.

Proof. Fix $T > 0$, and let $n > 0$.

• **STEP 1:** $\tilde{\mathbf{X}}^n \rightarrow \mathbf{X}$. Since $z \mapsto \tilde{X}^n(t, z)$ is non-decreasing with $\tilde{X}^n(t, 0) = x_0^n(t) \geq \bar{x}_0^n + v(R)t = \bar{x}_{\min} + v(R)t$ and $\tilde{X}^n(t, L) = \bar{x}_{\max} + v_{\max}t$, we have that

$$\begin{aligned} \text{TV}[\tilde{X}^n(t)] &\leq \bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R))t, \\ \|\tilde{X}^n(t)\|_{\mathbf{L}^\infty([0, L]; \mathbb{R})} &\leq \max\{|\bar{x}_{\min} + v(R)t|, |\bar{x}_{\max} + v_{\max}t|\}. \end{aligned}$$

Moreover, if $s < t$, then by (2.16b) and (3.5)

$$\begin{aligned} \int_0^L |\tilde{X}^n(t, z) - \tilde{X}^n(s, z)| dz &= \sum_{i=0}^{N_n-1} \ell_n |x_i^n(t) - x_i^n(s)| \\ &\leq \sum_{i=0}^{N_n-1} \ell_n \int_s^t |v(y_i^n(\tau))| d\tau \leq \max\{|v_{\max}|, |v(R)|\} L (t - s). \end{aligned}$$

Therefore, by applying Helly's theorem in the form [11, Theorem 2.4], up to a subsequence, $(\tilde{X}^n)_{n \in \mathbb{N}}$ converges in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L]; \mathbb{R})$ to a function X which is right continuous w.r.t. z and satisfying (3.8). Finally, since $\tilde{X}^{n+1}(t, z) \leq \tilde{X}^n(t, z)$ for all $t \geq 0$ and $z \in [0, L]$, the whole sequence $(\tilde{X}^n)_{n \in \mathbb{N}}$ converges to X and a.e. on $[0, +\infty[\times [0, L]$.

• **STEP 2:** $\hat{\mathbf{X}}^n \rightarrow \mathbf{X}$. By definition, see (2.17), (3.3) and (3.5), we have for all $t \in [0, T]$

$$\begin{aligned} \int_0^L |\hat{X}^n(t, z) - \tilde{X}^n(t, z)| dz &= \sum_{i=0}^{N_n-1} y_i^n(t)^{-1} \int_{i\ell_n}^{(i+1)\ell_n} [z - i\ell_n] dz \\ &= \frac{\ell_n}{2} \sum_{i=0}^{N_n-1} [x_{i+1}^n(t) - x_i^n(t)] = \frac{\ell_n}{2} [x_{N_n}^n(t) - x_0^n(t)] \\ &\leq \frac{\ell_n}{2} [\bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R))T], \end{aligned}$$

and the proof is complete as $(\tilde{X}^n)_{n \in \mathbb{N}}$ converges to X in view of STEP 1. \diamond

In the next lemma we prove that X inherits the maximum principle property satisfied by \tilde{X}^n proven in Lemma 2.1.

Lemma 3.1. *For all $t \geq 0$ and for a.e. $z_1, z_2 \in [0, L]$ with $z_1 < z_2$ we have*

$$\frac{z_2 - z_1}{R} \leq X(t, z_2) - X(t, z_1) \leq \bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R))t. \quad (3.9)$$

Proof. The upper bound is obvious. Take $0 \leq z_1 < z_2 \leq L$. For $n > 0$ sufficiently large, we can take $i, j \in \{0, 1, \dots, N_n\}$ such that $i < j$, $i \ell_n \leq z_1 < (i+1) \ell_n$ and $\ell_n j \leq z_2 < \ell_n (j+1)$. By (3.5) and Lemma 2.1 we have

$$\begin{aligned} \frac{\tilde{X}^n(t, z_2) - \tilde{X}^n(t, z_1)}{z_2 - z_1} &\geq \frac{x_j^n(t) - x_i^n(t)}{(j+1)\ell_n - i\ell_n} \geq \frac{(j-i)\ell_n}{R((j+1)\ell_n - i\ell_n)} \\ &= \frac{1}{R} - \frac{1}{R(j-i+1)} \geq \frac{1}{R} - \frac{1}{R((z_2 \ell_n^{-1} - 1) - z_1 \ell_n^{-1} + 1)} = \frac{1}{R} - \frac{\ell_n}{R(z_2 - z_1)}. \end{aligned}$$

By letting n go to infinity in the above estimate we conclude the proof. Indeed, as n goes to infinity we have that $\ell_n/[R(z_2 - z_1)]$ converges to zero and $(\tilde{X}^n)_{n \in \mathbb{N}}$ converges to X a.e. on $[0, +\infty[\times [0, L]$ in view of Proposition 3.1. \diamond

Proposition 3.2 (Definition of F). $(\hat{F}^n)_{n \in \mathbb{N}}$ and $(\tilde{F}^n)_{n \in \mathbb{N}}$ converge to $F \doteq \mathcal{F}[X]$ in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times \mathbb{R}; [0, L])$. Moreover, $(\tilde{F}^n)_{n \in \mathbb{N}}$ converges to F a.e. on $[0, +\infty[\times \mathbb{R}$.

Proof. We first observe that by Lemma 3.1 for any fixed $t \geq 0$, the map $z \mapsto X(t, z)$ is strictly increasing and for all $z \in [0, L]$

$$\frac{z}{R} + \bar{x}_{\min} + v(R)t \leq X(t, z) \leq \bar{x}_{\max} + v_{\max}t - \frac{L-z}{R}.$$

Thus, F is well defined. The convergence of $(\hat{F}^n)_{n \in \mathbb{N}}$ and $(\tilde{F}^n)_{n \in \mathbb{N}}$ to F follows from the basic property (2.14) of the scaled Wasserstein distance and from Proposition 3.1. Indeed, for any $T > 0$ we have

$$\begin{aligned} \lim_{n \rightarrow +\infty} \left\| \hat{F}^n - F \right\|_{\mathbf{L}^1([0, T] \times \mathbb{R}; \mathbb{R})} &= \lim_{n \rightarrow +\infty} \left\| \hat{X}^n - X \right\|_{\mathbf{L}^1([0, T] \times [0, L]; \mathbb{R})} = 0, \\ \lim_{n \rightarrow +\infty} \left\| \tilde{F}^n - F \right\|_{\mathbf{L}^1([0, T] \times \mathbb{R}; \mathbb{R})} &= \lim_{n \rightarrow +\infty} \left\| \tilde{X}^n - X \right\|_{\mathbf{L}^1([0, T] \times [0, L]; \mathbb{R})} = 0. \end{aligned}$$

Finally, $(\tilde{F}^n)_{n \in \mathbb{N}}$ converges to F a.e. on $[0, +\infty[\times \mathbb{R}$ because $\tilde{F}^{n+1}(t, x) \leq \tilde{F}^n(t, x)$ for all $t \geq 0$ and $x \in \mathbb{R}$. \diamond

Lemma 3.2. For all $t \geq 0$ and for a.e. $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$ we have

$$0 \leq F(t, x_2) - F(t, x_1) \leq R(x_2 - x_1). \quad (3.10)$$

Proof. Fix $x_1 < x_2$ and denote $z_1 = F(t, x_1) \leq z_2 = F(t, x_2)$. Since the lower bound is obvious, it is sufficient to prove that

$$z_2 - z_1 \leq R(x_2 - x_1).$$

If $z_1 = z_2$, then there is nothing to prove. Assume therefore that $z_1 \neq z_2$ and fix $\eta \in]0, z_2 - z_1[$. By definition, $X(t, z) = \mathcal{X}[F](t, z) = \inf\{x \in \mathbb{R} : F(t, x) > z\}$. Since $F(t, x_2) = z_2 > z_2 - \eta$, we have that $X(t, z_2 - \eta) \leq x_2$. Moreover, $X(t, z_1) \geq x_1$ because $z \mapsto X(t, z)$ is strictly increasing and right continuous. Therefore, by Lemma 3.1 we have

$$x_2 - x_1 \geq X(t, z_2 - \eta) - X(t, z_1) \geq \frac{z_2 - \eta - z_1}{R}.$$

Since $\eta > 0$ is arbitrary, we have $z_2 - z_1 \leq R(x_2 - x_1)$. \diamond

Proposition 3.3 (Definition of ρ). For any $t \geq 0$, let $\rho(t)$ be the distributional derivative of $x \mapsto F(t, x)$, with F defined in Lemma 3.1. Then:

- $\rho(t, \cdot) \in \mathcal{M}_L$ for all $t \geq 0$,
- $0 \leq \rho(t, x) \leq R$ for a.e. $t \geq 0$ and $x \in \mathbb{R}$,
- $(\hat{\rho}^n)_{n \in \mathbb{N}}$ and $(\tilde{\rho}^n)_{n \in \mathbb{N}}$ converge to ρ in the topology of $\mathbf{L}_{\text{loc}}^1([0, +\infty[; d_{L,1})$,

Proof. For any fixed $t \geq 0$, by Lemma 3.2 we have that $x \mapsto F(t, x)$ is a Lipschitz function with $\text{Lip}(F(t)) \leq R$. Hence its weak derivative $\rho(t)$ is well defined in the space of distributions and is essentially bounded with $\|\rho(t)\|_{\mathbf{L}^\infty(\mathbb{R})} \leq R$. Moreover, $x \mapsto F(t, x)$ is non-decreasing, and therefore $\rho(t) \geq 0$ a.e. in \mathbb{R} . By Proposition 3.1 and (2.14) we easily obtain that for any $T > 0$

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_0^T d_{L,1}(\hat{\rho}^n(t), \rho(t)) dt &= \lim_{n \rightarrow +\infty} \int_0^T \int_0^L |\hat{X}^n(t, z) - X(t, z)| dz dt = 0, \\ \lim_{n \rightarrow +\infty} \int_0^T d_{L,1}(\tilde{\rho}^n(t), \rho(t)) dt &= \lim_{n \rightarrow +\infty} \int_0^T \int_0^L |\tilde{X}^n(t, z) - X(t, z)| dz dt = 0. \end{aligned}$$

Thus, ρ satisfies also the last condition and $\rho(t) \in \mathcal{M}_L$. \diamond

Lemma 3.3 (Definition of $\check{\rho}$). *There exists $\check{\rho}$ in $\mathbf{L}^\infty([0, +\infty[\times [0, L]; \mathbb{R})$ such that, up to a subsequence, $(\check{\rho}^n)_{n \in \mathbb{N}}$ converges weakly- $*$ in $\mathbf{L}^\infty([0, +\infty[\times [0, L]; \mathbb{R})$ to $\check{\rho}$.*

Proof. It is sufficient to observe that for any $n > 0$ we have $0 \leq \check{\rho}^n \leq R$ a.e. on $[0, +\infty[\times [0, L]$ because, by Lemma 2.1, $\|y_i^n\|_{\mathbf{L}^\infty([0, +\infty[; \mathbb{R})} \leq R$. \diamond

We conclude this subsection by checking that the scheme is consistent with the prescribed initial condition in the limit.

Proposition 3.4. *The sequences $(\check{\rho}^n|_{t=0})_{n \in \mathbb{N}}$ and $(\hat{\rho}^n|_{t=0})_{n \in \mathbb{N}}$ both converge to $\bar{\rho}$ in the $d_{L,1}$ -Wasserstein distance.*

Proof. By definitions (2.19) and (2.20) we have that

$$\hat{\rho}^n(0, x) = \sum_{i=0}^{N_n-1} \bar{y}_i^n \chi_{[\bar{x}_i^n, \bar{x}_{i+1}^n[}(x), \quad \tilde{\rho}^n(0, x) = \ell_n \sum_{i=0}^{N_n-1} \delta_{\bar{x}_i^n}(x).$$

Therefore $F_{\hat{\rho}^n}|_{t=0} = \hat{F}^n|_{t=0}$, $F_{\tilde{\rho}^n}|_{t=0} = \tilde{F}^n|_{t=0}$ and by (2.14), (2.18c) we have

$$\begin{aligned} d_{L,1}(\tilde{\rho}^n|_{t=0}, \hat{\rho}^n|_{t=0}) &= \left\| \tilde{F}^n|_{t=0} - \hat{F}^n|_{t=0} \right\|_{\mathbf{L}^1(\mathbb{R}; \mathbb{R})} \\ &= \sum_{i=0}^{N_n-2} \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \left[\ell_n - \bar{y}_i^n [x - \bar{x}_i^n] \right] dx = \ell_n \sum_{i=0}^{N_n-2} \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \frac{\bar{x}_{i+1}^n - x}{\bar{x}_{i+1}^n - \bar{x}_i^n} dx \\ &\leq \ell_n [\bar{x}_{\max} - \bar{x}_{\min}]. \end{aligned}$$

Hence, it is sufficient to prove that $(\tilde{\rho}^n|_{t=0})_{n \in \mathbb{N}}$ converges to $\bar{\rho}$ in the $d_{L,1}$ -Wasserstein distance. By (2.15) we have that

$$\begin{aligned} d_{L,1}(\tilde{\rho}^n|_{t=0}, \bar{\rho}) &= \left\| \tilde{F}^n|_{t=0} - F_{\bar{\rho}} \right\|_{\mathbf{L}^1(\mathbb{R}; \mathbb{R})} = \sum_{i=0}^{N_n-2} \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \left[\ell_n (i+1) - \int_{-\infty}^x \bar{\rho}(y) dy \right] \\ &= \sum_{i=0}^{N_n-2} \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \left[\ell_n - \int_{\bar{x}_i^n}^x \bar{\rho}(y) dy \right] dx \leq \ell_n [\bar{x}_{\max} - \bar{x}_{\min}] \end{aligned}$$

and this concludes the proof. \diamond

3.2. BV estimates and discrete Oleinik condition. Let us sum up what we have proven so far. The family of empirical measures $(\tilde{\rho}^n)_{n \in \mathbb{N}}$ converges in the scaled 1-Wasserstein sense to a limit ρ belonging to $\mathbf{L}^\infty([0, +\infty[; \mathcal{M}_L)$ and such that $0 \leq \rho \leq R$ almost everywhere. Moreover, the empirical measure $\tilde{\rho}^n$ has a pseudo-inverse distribution function \tilde{X}^n satisfying the PDE

$$\tilde{X}_t^n(t, z) = v(\tilde{\rho}^n(t, z)), \quad (t, z) \in [0, +\infty[\times [0, L],$$

with the family $(\tilde{\rho}^n)_{n \in \mathbb{N}}$ being weakly-* compact in $\mathbf{L}^\infty([0, +\infty[\times [0, L]; [0, 1])$. The Wasserstein topology is a proper tool to pass to the limit the time derivative term in the above PDE, as this term is linear. But on the other hand, the weak-* topology is too weak to pass to the limit $v(\tilde{\rho}^n)$ for a general nonlinear v . Moreover, there is the additional difficulty of having to check that the two limits are related in some sense.

A typical way to overcome the difficulty stated above is to provide a **BV** estimate for the approximating sequence $(\tilde{\rho}^n)_{n \in \mathbb{N}}$. We tackle this task in two ways. First of all, we perform a direct estimate of the total variation of $\tilde{\rho}^n$, and prove that such a quantity decreases in time, and is therefore uniformly bounded provided the initial datum $\bar{\rho}$ is **BV**. However, this result is only partly satisfactory, as it is well known that the solution ρ to (2.1a) is **BV** even for an initial datum in $\mathbf{L}^1 \cap \mathbf{L}^\infty$. We shall therefore prove that a uniform **BV** estimate of $\tilde{\rho}^n$ is available for an initial datum in $\mathbf{L}^1 \cap \mathbf{L}^\infty$ provided the additional property (V3) of v is prescribed. The latter task is performed by a one-sided estimate of the difference quotients of $\tilde{\rho}^n$, in the spirit of a discrete version of the classical Oleinik-type condition (2.5), which can be considered as the main technical achievement of this paper.

We start with the following proposition.

Proposition 3.5 (**BV** contractivity for **BV** initial data). *Assume v satisfies (V1) and (V2). If $\bar{\rho}$ satisfies (InBV), then for any $n \in \mathbb{N}$*

$$\mathrm{TV}[\tilde{\rho}^n(t)] = \mathrm{TV}[\tilde{\rho}^n(t)] \leq \mathrm{TV}[\bar{\rho}] \quad \text{for all } t \geq 0.$$

Proof. For notational simplicity, we shall omit the dependence on t and n whenever not necessary. By construction, see (2.17) and (2.19), we have that

$$\begin{aligned} \mathrm{TV}[\tilde{\rho}(0)] &= \bar{y}_0 + \bar{y}_{N-1} + \sum_{i=0}^{N-2} |\bar{y}_i - \bar{y}_{i+1}| \\ &= \int_{\bar{x}_{\min}}^{\bar{x}_1} \bar{\rho}(y) dy + \int_{\bar{x}_{N-1}}^{\bar{x}_{\max}} \bar{\rho}(y) dy + \sum_{i=0}^{N-2} \left| \int_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{\rho}(y) dy - \int_{\bar{x}_{i+2}}^{\bar{x}_{i+1}} \bar{\rho}(y) dy \right| \leq \mathrm{TV}[\bar{\rho}]. \end{aligned}$$

Moreover

$$\begin{aligned} \frac{d}{dt} \mathrm{TV}[\tilde{\rho}(t)] &= \frac{d}{dt} \left[y_0 + y_{N-1} + \sum_{i=0}^{N-2} |y_i - y_{i+1}| \right] = \dot{y}_0 + \dot{y}_{N-1} \\ &+ \sum_{i=0}^{N-2} \operatorname{sgn}[y_i - y_{i+1}] [\dot{y}_i - \dot{y}_{i+1}] = \left[1 + \operatorname{sgn}[y_0 - y_1] \right] \dot{y}_0 \\ &+ \left[1 - \operatorname{sgn}[y_{N-2} - y_{N-1}] \right] \dot{y}_{N-1} + \sum_{i=1}^{N-2} \left[\operatorname{sgn}[y_i - y_{i+1}] - \operatorname{sgn}[y_{i-1} - y_i] \right] \dot{y}_i. \end{aligned}$$

We claim that the latter right hand side above is ≤ 0 . Indeed, assumptions (V1) and (V2) together with (2.18) imply that the following quantities are not positive

$$\left[1 + \operatorname{sgn}[y_0 - y_1] \right] \dot{y}_0 = - \left[1 + \operatorname{sgn}[y_0 - y_1] \right] \frac{y_0^2}{\ell} [v(y_1) - v(y_0)],$$

$$\begin{aligned}
\left[1 - \operatorname{sgn}[y_{N-2} - y_{N-1}]\right] \dot{y}_{N-1} &= - \left[1 - \operatorname{sgn}[y_{N-2} - y_{N-1}]\right] \times \\
&\quad \times \frac{y_{N-1}^2}{\ell} [v_{\max} - v(y_{N-1})], \\
\left[\operatorname{sgn}[y_i - y_{i+1}] - \operatorname{sgn}[y_{i-1} - y_i]\right] \dot{y}_i &= - \left[\operatorname{sgn}[y_i - y_{i+1}] - \operatorname{sgn}[y_{i-1} - y_i]\right] \times \\
&\quad \times \frac{y_i^2}{\ell} [v(y_{i+1}) - v(y_i)].
\end{aligned}$$

Therefore, $\operatorname{TV}[\hat{\rho}(t)] \leq \operatorname{TV}[\hat{\rho}]$ for all $t \geq 0$. Finally, since $\tilde{\rho}^n = \hat{\rho} \circ \hat{X}$, $\tilde{\rho}^n$ is piecewise constant and it has on $i\ell$ the same traces as $\hat{\rho}$ on x_i , the statement for $\tilde{\rho}^n$ follows easily. \diamond

We now perform our discrete Oleinik-type condition, which holds for general initial data in $\mathbf{L}^1 \cap \mathbf{L}^\infty$.

Lemma 3.4 (Discrete Oleinik-type condition). *Assume v satisfies (V1), (V2), and (V3), and let $\bar{\rho}$ satisfy (In). Then, for any $i = 0, \dots, N_n - 2$ we have*

$$t y_i^n(t) [v(y_{i+1}^n(t)) - v(y_i^n(t))] \leq \ell_n \quad \text{for all } t \geq 0. \quad (3.11)$$

Proof. For notational simplicity, we shall omit the dependence on t and n whenever not necessary. Let

$$\begin{aligned}
z_i &\doteq t y_i [v(y_{i+1}) - v(y_i)], & i = 0, \dots, N-2, \\
z_{N-1} &\doteq t y_{N-1} [v_{\max} - v(y_{N-1})].
\end{aligned}$$

• **STEP 0:** $z_{N-1} \leq \ell$. By (2.18a) and (V1)

$$\begin{aligned}
\dot{z}_{N-1} &= y_{N-1} [v_{\max} - v(y_{N-1})] + t \dot{y}_{N-1} [v_{\max} - v(y_{N-1})] - t y_{N-1} v'(y_{N-1}) \dot{y}_{N-1} \\
&= y_{N-1} [v_{\max} - v(y_{N-1})] - \frac{t y_{N-1}^2}{\ell} [v_{\max} - v(y_{N-1})]^2 \\
&\quad + \frac{t v'(y_{N-1}) y_{N-1}^3}{\ell} [v_{\max} - v(y_{N-1})] \leq y_{N-1} [v_{\max} - v(y_{N-1})] \left[1 - \frac{z_{N-1}}{\ell}\right].
\end{aligned}$$

Since $z_{N-1}(0) = 0$, from the above estimate we get $z_{N-1}(t) \leq \ell$ for all $t \geq 0$. Indeed, assume by contradiction that there exist $t_1 < t_2$ such that $z_{N-1}(t_1) = \ell$ and $z_{N-1}(t) > \ell$ for all $t \in]t_1, t_2[$. The above estimate implies for $t \in]t_1, t_2[$

$$\begin{aligned}
z_{N-1}(t) &= z_{N-1}(t_1) + \int_{t_1}^t y_{N-1}(s) [v_{\max} - v(y_{N-1}(s))] \left[1 - \frac{z_{N-1}(s)}{\ell}\right] ds \\
&\leq z_{N-1}(t_1) = \ell,
\end{aligned}$$

which gives a contradiction.

• **STEP 1:** $z_{i+1} \leq \ell \Rightarrow z_i \leq \ell$. Let $i \in \{0, \dots, N-3\}$ and assume $z_{i+1} \leq \ell$. From (2.18b) and (V1) we get

$$\begin{aligned}
\dot{z}_i &= y_i [v(y_{i+1}) - v(y_i)] + t \dot{y}_i [v(y_{i+1}) - v(y_i)] + t y_i [v'(y_{i+1}) \dot{y}_{i+1} - v'(y_i) \dot{y}_i] \\
&= y_i [v(y_{i+1}) - v(y_i)] - \frac{t y_i^2}{\ell} [v(y_{i+1}) - v(y_i)]^2 \\
&\quad + t y_i \left[-\frac{v'(y_{i+1}) y_{i+1}^2}{\ell} [v(y_{i+2}) - v(y_{i+1})] + \frac{v'(y_i) y_i^2}{\ell} [v(y_{i+1}) - v(y_i)] \right] \\
&= y_i [v(y_{i+1}) - v(y_i)] - \frac{y_i}{\ell} [v(y_{i+1}) - v(y_i)] z_i - \frac{v'(y_{i+1}) y_i y_{i+1}}{\ell} z_{i+1} + \frac{v'(y_i) y_i^2}{\ell} z_i.
\end{aligned}$$

Since $\operatorname{sgn}_+[z_i] = \operatorname{sgn}_+[v(y_{i+1}) - v(y_i)] = \operatorname{sgn}_+[y_i - y_{i+1}]$ for all $t > 0$, from the assumption on z_{i+1} we easily obtain

$$\begin{aligned} \frac{d}{dt}[z_i]_+ &= y_i [v(y_{i+1}) - v(y_i)]_+ - \frac{y_i}{\ell} [v(y_{i+1}) - v(y_i)]_+ [z_i]_+ \\ &\quad - \frac{v'(y_{i+1}) y_i y_{i+1}}{\ell} \operatorname{sgn}_+[z_i] z_{i+1} + \frac{v'(y_i) y_i^2}{\ell} [z_i]_+ \\ &\leq y_i [v(y_{i+1}) - v(y_i)]_+ \left[1 - \frac{[z_i]_+}{\ell} \right] - v'(y_{i+1}) y_i y_{i+1} \operatorname{sgn}_+[z_i] + \frac{v'(y_i) y_i^2}{\ell} [z_i]_+. \end{aligned}$$

Condition (V3) prescribes that the function $y \mapsto y v'(y)$ is non-increasing, which gives

$$\begin{aligned} \frac{d}{dt}[z_i]_+ &\leq y_i [v(y_{i+1}) - v(y_i)]_+ \left[1 - \frac{[z_i]_+}{\ell} \right] - v'(y_i) y_i^2 \operatorname{sgn}_+[z_i] + \frac{v'(y_i) y_i^2}{\ell} [z_i]_+ \\ &= y_i \left[[v(y_{i+1}) - v(y_i)]_+ - v'(y_i) y_i \right] \left[1 - \frac{[z_i]_+}{\ell} \right]. \end{aligned}$$

Now, as $v' \leq 0$, and since $z_i(0) = 0$, by a similar comparison argument as the one at the end of Step 1 we get that $z_i(t)_+ \leq \ell$ for all $t \geq 0$.

• **STEP 2:** $z_{N-2} \leq \ell$. From analogous computations as in previous step, we get

$$\begin{aligned} \frac{d}{dt}[z_{N-2}]_+ &= y_{N-2} [v(y_{N-1}) - v(y_{N-2})]_+ - \frac{y_{N-2}}{\ell} [v(y_{N-1}) - v(y_{N-2})]_+ [z_{N-2}]_+ \\ &\quad - \frac{v'(y_{N-1}) y_{N-2} y_{N-1}}{\ell} \operatorname{sgn}_+[z_{N-2}] z_{N-1} + \frac{v'(y_{N-2}) y_{N-2}^2}{\ell} [z_{N-2}]_+, \end{aligned}$$

and we can use the monotonicity of $y \mapsto y v'(y)$ and Step 0 to get

$$\begin{aligned} \frac{d}{dt}[z_{N-2}]_+ &\leq y_{N-2} [v(y_{N-1}) - v(y_{N-2})]_+ \left[1 - \frac{[z_{N-2}]_+}{\ell} \right] \\ &\quad - v'(y_{N-2}) y_{N-2}^2 \operatorname{sgn}_+[z_{N-2}] + \frac{v'(y_{N-2}) y_{N-2}^2}{\ell} [z_{N-2}]_+ \\ &= y_{N-2} \left[[v(y_{N-1}) - v(y_{N-2})]_+ - v'(y_{N-2}) y_{N-2} \right] \left[1 - \frac{[z_{N-2}]_+}{\ell} \right]. \end{aligned}$$

Again, $v' \leq 0$ and $z_{N-2}(0) = 0$ imply that $z_{N-2}(t)_+ \leq \ell$ for all $t \geq 0$.

• **CONCLUSION.** The estimate (3.11) is proven recursively: Step 2 provides the first step with $i = N - 2$, whereas Step 1 proves that the estimate holds for all $i \in \{0, \dots, N - 3\}$. \diamond

Corollary 3.1. *Assume v satisfies (V1), (V2), and (V3), and let $\bar{\rho}$ satisfy (In). Then, for any $i \in \{0, \dots, N - 2\}$ we have*

$$v(\hat{\rho}^n(t, x_i^n(t))) - v(\hat{\rho}^n(t, x_{i+1}^n(t))) \leq \frac{x_{i+1}^n(t) - x_i^n(t)}{t} \quad \text{for all } t > 0. \quad (3.12)$$

Proof. The statement follows from Lemma 3.4 and the definitions of $\hat{\rho}^N$ and y_i . \diamond

In the following proposition we prove uniform bounds on the total variation of $v(\check{\rho}^n)$ and $v(\hat{\rho}^n)$. Let us emphasize that the regularising effect $\mathbf{L}^\infty \mapsto \mathbf{BV}$ implies that the \mathbf{BV} estimate eventually *blows up* as $t \searrow 0$.

Proposition 3.6 (Uniform \mathbf{BV} estimates for $v(\check{\rho}^n)$ and $v(\hat{\rho}^n)$). *Assume v satisfies the properties (V1), (V2), and (V3), and let $\bar{\rho}$ satisfy (In). Let $\delta > 0$. Then*

- (i) $(v(\hat{\rho}^n))_{n \in \mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty([\delta, +\infty[; \mathbf{BV}(\mathbb{R}; [v(R), v_{\max}]))$;
- (ii) $(v(\check{\rho}^n))_{n \in \mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty([\delta, +\infty[; \mathbf{BV}([0, L]; [v(R), v_{\max}]))$.

More precisely, for any $n \in \mathbb{N}$

$$\mathrm{TV} [v(\hat{\rho}^n(t))] = \mathrm{TV} [v(\check{\rho}^n(t))] \leq C_\delta \quad \text{for all } t \geq \delta,$$

where $C_\delta \doteq [3(v_{\max} - v(R)) + 2 \frac{\bar{x}_{\max} - \bar{x}_{\min}}{\delta}]$.

Proof. For notational simplicity, we shall omit the dependence on n . We set

$$\hat{\sigma}(t, x) \doteq v(\hat{\rho}(t, x)) + \frac{1}{t} \sum_{i=0}^{N-1} x_i(t) \chi_{[x_i(t), x_{i+1}(t)]}(x) \quad \text{for all } x \in \mathbb{R}.$$

We claim that, for any fixed $t \geq 0$, the map $x \mapsto \hat{\sigma}(t, x)$ is a piecewise constant, non-decreasing function on $[x_0(t), x_N(t)]$. To see this, we first notice that the map $x \mapsto \hat{\sigma}(t, x)$ is constant on the interval $[x_i(t), x_{i+1}(t)]$, $i = 0, \dots, N-1$. On the other hand, $\hat{\sigma}(t)$ is non-decreasing on the potential discontinuity points $x_i(t)$, $i = 1, \dots, N-1$, in view of (3.12). Now, from (2.16) and the discrete maximum principle in Lemma 2.1 we know that for any $x \in [x_0(t), x_N(t)]$

$$\frac{\bar{x}_{\min}}{t} + 2v(R) \leq \hat{\sigma}(t, x) \leq v_{\max} + \frac{1}{t} [\bar{x}_{\max} + v_{\max} t] = 2v_{\max} + \frac{\bar{x}_{\max}}{t}.$$

Hence $\hat{\sigma}$ is uniformly bounded in $\mathbf{L}^\infty([\delta, +\infty[; \mathbf{BV}(\mathbb{R}; \mathbb{R}))$ with

$$\sup_{t \geq \delta} \mathrm{TV} [\hat{\sigma}(t)] \leq \left[2(v_{\max} - v(R)) + \frac{\bar{x}_{\max} - \bar{x}_{\min}}{\delta} \right].$$

Therefore, also $v(\hat{\rho})$ is uniformly bounded in $\mathbf{L}^\infty([\delta, +\infty[; \mathbf{BV}(\mathbb{R}; \mathbb{R}))$ because by triangular inequality

$$\begin{aligned} \mathrm{TV} [v(\hat{\rho}(t))] &\leq \mathrm{TV} [\hat{\sigma}(t)] + \mathrm{TV} \left[\frac{1}{t} \sum_{i=0}^{N-1} x_i(t) \chi_{[x_i(t), x_{i+1}(t)]} \right] \\ &= \mathrm{TV} [\hat{\sigma}(t)] + \frac{1}{t} [\bar{x}_{\max} - \bar{x}_{\min} + (v_{\max} - v(R)) t] \leq C_\delta, \end{aligned}$$

for $t \geq \delta$. Finally, since $\check{\rho}^n = \hat{\rho} \circ \hat{X}$, $\check{\rho}^n$ is piecewise constant and on $i\ell$ has the same traces as $\hat{\rho}$ on x_i , the statement for $\check{\rho}^n$ follows easily. \diamond

3.3. Time continuity and compactness.

Proposition 3.7 (Uniform \mathbf{L}^1 -continuity in time of $\check{\rho}^n$). *Under the assumptions of Proposition 3.6, for any $\delta > 0$ we have*

$$\int_0^L |\check{\rho}^n(t, z) - \check{\rho}^n(s, z)| \, dz \leq R^2 [C_\delta + (v_{\max} - v(R))] |t - s| \quad \text{for all } t, s \geq \delta,$$

with C_δ defined in Proposition 3.6.

Proof. By (2.18), we compute for $t > s > \delta$,

$$\begin{aligned} \int_0^L |\check{\rho}^n(t, z) - \check{\rho}^n(s, z)| \, dz &= \sum_{i=0}^{N_n-1} \ell_n |y_i^n(t) - y_i^n(s)| = \sum_{i=0}^{N_n-1} \ell_n \left| \int_s^t \dot{y}_i^n(\tau) \, d\tau \right| \\ &= \sum_{i=0}^{N_n-2} \left| \int_s^t y_i^n(\tau)^2 [v(y_{i+1}^n(\tau)) - v(y_i^n(\tau))] \, d\tau \right| + \int_s^t y_{N_n-1}^n(\tau)^2 [v_{\max} - v(y_{N_n-1}^n(\tau))] \, d\tau. \end{aligned}$$

Therefore, by Lemma 2.1

$$\int_0^L |\check{\rho}^n(t, z) - \check{\rho}^n(s, z)| \, dz$$

$$\begin{aligned}
&\leq \int_s^t \left[\sum_{i=0}^{N_n-2} y_i^n(\tau)^2 |v(y_{i+1}^n(\tau)) - v(y_i^n(\tau))| + y_{N_n-1}^n(\tau)^2 [v_{\max} - v(y_{N_n-1}^n(\tau))] \right] d\tau \\
&\leq R^2 \int_s^t \left[\sum_{i=0}^{N_n-2} |v(y_{i+1}^n(\tau)) - v(y_i^n(\tau))| + (v_{\max} - v(R)) \right] d\tau \\
&\leq R^2 \int_s^t \left[\text{TV}[v(\hat{\rho}^n(\tau))] + (v_{\max} - v(R)) \right] d\tau.
\end{aligned}$$

Then it is sufficient to apply the estimate in Proposition 3.6 to complete the proof. \diamond

Proposition 3.8 (Uniform Wasserstein time continuity of $\hat{\rho}^n$). *Assume v satisfies the properties (V1) and (V2), and let $\bar{\rho}$ satisfy the assumption (In). For any $n \in \mathbb{N}$ we have*

$$d_{L,1}(\hat{\rho}^n(t), \hat{\rho}^n(s)) \leq 2L \max\{|v_{\max}|, |v(R)|, [v_{\max} - v(R)]\} |t - s| \quad \text{for all } s, t \geq 0.$$

Proof. By (2.14), (2.17) and (2.16), we compute for any $t > s \geq 0$

$$\begin{aligned}
d_{L,1}(\hat{\rho}^n(t), \hat{\rho}^n(s)) &= \left\| \hat{X}^n(t) - \hat{X}^n(s) \right\|_{\mathbf{L}^1([0,L];\mathbb{R})} \\
&= \sum_{i=0}^{N_n-1} \int_{i\ell_n}^{(i+1)\ell_n} \left| \hat{X}^n(t, z) - \hat{X}^n(s, z) \right| dz \\
&= \sum_{i=0}^{N_n-1} \int_{i\ell_n}^{(i+1)\ell_n} \left| x_i^n(t) + \frac{z - i\ell_n}{y_i^n(t)} - x_i^n(s) - \frac{z - i\ell_n}{y_i^n(s)} \right| dz \\
&\leq \sum_{i=0}^{N_n-1} \ell_n |x_i^n(t) - x_i^n(s)| + \sum_{i=0}^{N_n-1} |y_i^n(t)^{-1} - y_i^n(s)^{-1}| \int_{i\ell_n}^{(i+1)\ell_n} (z - i\ell_n) dz \\
&\leq \sum_{i=0}^{N_n-1} \ell_n \int_s^t |v(y_i^n(\tau))| d\tau + \sum_{i=0}^{N_n-1} \frac{\ell_n^2}{2} \int_s^t \left| \frac{d}{d\tau} [y_i^n(\tau)^{-1}] \right| d\tau \\
&\leq L \max\{|v_{\max}|, |v(R)|\} (t - s) \\
&\quad + \frac{\ell_n}{2} \int_s^t \left[\sum_{i=0}^{N_n-2} |v(y_{i+1}^n(\tau)) - v(y_i^n(\tau))| + [v_{\max} - v(y_{N_n-1}^n(\tau))] \right] d\tau \\
&= L \max\{|v_{\max}|, |v(R)|\} (t - s) + L \int_s^t [v_{\max} - v(R)] d\tau \\
&\leq 2L \max\{|v_{\max}|, |v(R)|, [v_{\max} - v(R)]\} (t - s)
\end{aligned}$$

and this concludes the proof. \diamond

We now recall a generalization of Aubin-Lions lemma, which uses the Wasserstein distance as a replacement of a negative Sobolev norm, proven in [47, Theorem 2], which we present here in a version adapted to our case. In order to have the paper self-contained, we first recall the precise statement of [47, Theorem 2] (see also the adapted version in [21]).

Theorem 3.1 (Theorem 2 from [47]). *On a separable Banach space \mathbb{X} , let be given*

- (F) a normal coercive integrand $\mathfrak{F} : \mathbb{X} \rightarrow [0, +\infty]$, i.e., \mathfrak{F} is lower semi-continuous and its sublevels are relatively compact in \mathbb{X} ;
- (g) a pseudo-distance $\mathfrak{g} : \mathbb{X} \times \mathbb{X} \rightarrow [0, +\infty]$, i.e., \mathfrak{g} is lower semi-continuous, and if $\nu, \mu \in \mathbb{X}$ are such that $\mathfrak{g}(\nu, \mu) = 0$, $\mathfrak{F}[\nu] < +\infty$ and $\mathfrak{F}[\mu] < +\infty$, then $\nu = \mu$.

Let further U be a set of measurable functions $\nu :]0, T[\rightarrow \mathbb{X}$, with a fixed $T > 0$. Under the hypotheses that

$$\sup_{\nu \in U} \int_0^T \mathfrak{F}[\nu(t)] dt < +\infty \quad \text{and} \quad \lim_{h \downarrow 0} \left[\sup_{\nu \in U} \int_0^{T-h} \mathfrak{g}(\nu(t+h), \nu(t)) dt \right] = 0, \quad (3.13)$$

Then U is strongly relatively compact in $\mathbf{L}^1([0, T[; \mathbb{X})$.

Theorem 3.2 (Generalized Aubin-Lions lemma). *Let $T, L > 0$ and $I \subset \mathbb{R}$ be a bounded open convex interval. Assume $w : \mathbb{R} \rightarrow \mathbb{R}$ is a Lipschitz continuous and strictly monotone function. Let $(\rho^n)_{n \in \mathbb{N}}$ be a sequence in $\mathbf{L}^\infty([0, T[\times \mathbb{R}) \cap \mathcal{M}_L$ such that*

- (1) $\rho^n :]0, T[\rightarrow \mathbf{L}^1(\mathbb{R})$ is measurable for all $n \in \mathbb{N}$;
- (2) $\text{spt}(\rho^n(t)) \subseteq I$ for all $t \in]0, T[$ and $n \in \mathbb{N}$;
- (3) $\sup_{n \in \mathbb{N}} \int_0^T [\|w(\rho^n(t))\|_{\mathbf{L}^1(I)} + \text{TV}[w(\rho^n(t))]] dt < +\infty$;
- (4) there exists a constant C depending only on T such that $d_{L,1}(\rho^n(s), \rho^n(t)) \leq C|t - s|$ for all $s, t \in]0, T[$ and $n \in \mathbb{N}$.

Then, $(\rho^n)_{n \in \mathbb{N}}$ is strongly relatively compact in $\mathbf{L}^1([0, T[\times \mathbb{R})$.

Proof. We want to use Theorem 3.1 with

$$\begin{aligned} \mathbb{X} &\doteq \mathbf{L}^1(I), & U &\doteq (\rho^n)_{n \in \mathbb{N}}, \\ \mathfrak{F}[\nu] &\doteq \|w(\nu)\|_{\mathbf{L}^1(I)} + \text{TV}[w(\nu)], & \mathfrak{g}(\nu, \mu) &\doteq \begin{cases} d_{L,1}(\nu, \mu) & \text{if } \nu, \mu \in \mathcal{M}_L, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

We first have to prove that \mathfrak{F} , \mathfrak{g} and U satisfy the corresponding hypotheses in Theorem 3.1.

(F) Assume that $(\nu^n)_{n \in \mathbb{N}}$ converges to ν strongly in $\mathbf{L}^1(I)$. Since w is Lipschitz continuous, $(w(\nu^n))_{n \in \mathbb{N}}$ converges to $w(\nu)$ strongly in $\mathbf{L}^1(I)$. Hence, for the lower semi-continuity of the total variation w.r.t. the \mathbf{L}^1 -norm, see [24, Theorem 1 on page 172], we have that $\text{TV}[w(\nu)] \leq \liminf_{n \rightarrow +\infty} \text{TV}[w(\nu^n)]$. Thus $\mathfrak{F}[\nu] \leq \liminf_{n \rightarrow +\infty} \mathfrak{F}[\nu^n]$ and \mathfrak{F} is l.s.c. in \mathbb{X} . Finally, consider a sequence $(\nu^n)_{n \in \mathbb{N}}$ belonging to a sublevel of \mathfrak{F} , namely $\sup_{n \in \mathbb{N}} \mathfrak{F}[\nu^n] < +\infty$. For the compactness of \mathbf{BV} in \mathbf{L}^1 on bounded open convex intervals and for basic properties of the \mathbf{L}^1 -convergence, see [24, Theorem 4 on page 176 and Theorem 5 on page 21], up to a subsequence $(w(\nu^n))_{n \in \mathbb{N}}$ converges to \bar{w} in \mathbf{L}^1 and a.e. on I . Since w is continuous and strictly monotone, $(\nu^n)_{n \in \mathbb{N}}$ is uniformly bounded in \mathbf{L}^∞ (consequence of the uniform bound on the total variation) and converges to $\bar{\nu} \doteq w^{-1}(\bar{w})$ a.e. on I and therefore, by the Lebesgue dominated convergence theorem, the convergence is also in \mathbf{L}^1 .

(g) Proceeding as before and applying lower semi-continuity of the 1-Wasserstein distance w.r.t. the \mathbf{L}^1 -norm give that \mathfrak{g} is l.s.c. in $\mathbb{X} \times \mathbb{X}$. Finally, if $\mathfrak{F}[\nu] < +\infty$, $\mathfrak{F}[\mu] < +\infty$ and $\mathfrak{g}(\mu, \nu) = 0$, then $w(\mu), w(\nu)$ are in \mathbf{BV} , $\nu, \mu \in \mathcal{M}_L$, and $d_{L,1}(\mu, \nu) = 0$. Hence we have $\mu = \nu$.

(U) Conditions in (3.13) follow directly from the hypotheses (3) and (4).

Hence we can apply Theorem 3.1 and obtain the strong compactness in $\mathbf{L}^1([0, T[\times I)$. Finally, recalling the hypothesis 2 concludes the proof. \diamond

3.4. Convergence to entropy solutions. In the next proposition we collect the previous compactness results to get strong convergence.

Proposition 3.9. *Let $\check{\rho}$ be defined as in Lemma 3.3 and ρ as in Proposition 3.3. Under the assumptions in Theorem 2.3 we have that*

- (i) the sequence $(\check{\rho}^n)_{n \in \mathbb{N}}$ converges up to a subsequence to $\check{\rho}$ almost everywhere and strongly in $\mathbf{L}_{\text{loc}}^1$ on $]0, +\infty[\times [0, L]$;
- (ii) the sequence $(\hat{\rho}^n)_{n \in \mathbb{N}}$ converges up to a subsequence to ρ almost everywhere and strongly in $\mathbf{L}_{\text{loc}}^1$ on $]0, +\infty[\times \mathbb{R}$;

(iii) if $\bar{\rho}$ satisfies also (InBV), then the sequence $(\check{\rho}^n)_{n \in \mathbb{N}}$ converges up to a subsequence to $\check{\rho}$ strongly in $\mathbf{L}_{\text{loc}}^1$ on $[0, +\infty[\times [0, L]$.

Proof. We already know from Proposition 3.3 that both $(\hat{\rho}^n)_{n \in \mathbb{N}}$ and $(\check{\rho}^n)_{n \in \mathbb{N}}$, defined respectively by (2.19) and (2.20), converge in the topology of $\mathbf{L}_{\text{loc}}^1([0, +\infty[; d_{L,1})$ to the density $\rho \in \mathbf{L}^\infty([0, +\infty[; \mathcal{M}_L)$ with $0 \leq \rho \leq R$. From Proposition 3.1 we know that both $(\hat{X}^n)_{n \in \mathbb{N}}$ and $(\tilde{X}^n)_{n \in \mathbb{N}}$, defined respectively by (3.3) and (3.5), converge strongly in $\mathbf{L}_{\text{loc}}^1([0, +\infty[\times [0, L])$ to $X \in \mathbf{L}^\infty([0, +\infty[\times [0, L])$, the pseudo-inverse of F , the cumulative distribution of ρ . Finally, from Lemma 3.3 we know that $(\check{\rho}^n)_{n \in \mathbb{N}}$, defined by (3.6), converges up to a subsequence to $\check{\rho}$ weakly-* in $\mathbf{L}^\infty([0, +\infty[\times [0, L])$.

• **STEP 1.** *Strong convergence of $(\check{\rho}^n)_{n \in \mathbb{N}}$ for general initial datum in $\mathcal{M}_L \cap \mathbf{L}^\infty$.*

Let $\bar{\rho}$ satisfy (In). For any fixed $\delta > 0$, we know from Proposition 3.6 that $(v(\check{\rho}^n))_{n \in \mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty([\delta, +\infty[; \mathbf{BV}([0, L]; [v(R), v_{\max}]))$. Furthermore, from Proposition 3.7 we easily obtain that

$$\int_0^L |v(\check{\rho}^n(t, z)) - v(\check{\rho}^n(s, z))| dz \leq \text{Lip}(v) R^2 [C_\delta + (v_{\max} - v(R))] |t - s|,$$

for all $t, s \geq \delta$. Therefore, we can apply Helly's theorem in the form [11, Theorem 2.4] to get that $(v(\check{\rho}^n))_{n \in \mathbb{N}}$ is strongly compact in $\mathbf{L}_{\text{loc}}^1([\delta, +\infty[\times [0, L])$. Hence, by the monotonicity of v , up to a subsequence $(\check{\rho}^n)_{n \in \mathbb{N}}$ converges strongly in $\mathbf{L}_{\text{loc}}^1$ and a.e. on $[\delta, +\infty[\times [0, L]$ to $\check{\rho}$. Finally, since $\delta > 0$ is arbitrary, the proof of (i) is complete.

• **STEP 2.** *Strong convergence of $(\hat{\rho}^n)_{n \in \mathbb{N}}$ for general initial datum in $\mathcal{M}_L \cap \mathbf{L}^\infty$.*

Let $\bar{\rho}$ satisfy (In) and fix $T, \delta > 0$ with $\delta < T$. We want to prove that $(\rho^n)_{n \in \mathbb{N}}$ with $\rho^n(t, x) \doteq \hat{\rho}^n(t + \delta, x)$ satisfies the hypotheses of Theorem 3.2 with

$$I =]\bar{x}_{\min} + v(R)(T + \delta) - 1, \bar{x}_{\max} + v_{\max}(T + \delta) + 1[$$

and $w = v$. The hypotheses 1 and 2 are satisfied because by (2.19) we have that $\|\hat{\rho}^n(t)\|_{\mathbf{L}^1(\mathbb{R})} = L$ and $\text{spt}(\hat{\rho}^n(t)) \subset I$ for all $t \in [0, T + \delta]$. By Proposition 3.6, the hypothesis 3 holds true because

$$\int_\delta^{T+\delta} \left[\|v(\hat{\rho}^n(t))\|_{\mathbf{L}^1(I; \mathbb{R})} + \text{TV}[v(\hat{\rho}^n(t))] \right] dt \leq [\max\{|v_{\max}|, |v(R)|\} |I| + C_\delta] T.$$

Finally, the hypothesis 4 follows directly from Proposition 3.8. Hence, we can apply Theorem 3.2 to obtain that $(\hat{\rho}^n)_{n \in \mathbb{N}}$ is strongly compact in $\mathbf{L}^1([\delta, T[\times \mathbb{R}; \mathbb{R})$. By the uniqueness of the limit in the $\mathbf{L}^1([\delta, T[; d_{L,1})$ topology, up to a subsequence $(\hat{\rho}^n)_{n \in \mathbb{N}}$ converges strongly in \mathbf{L}^1 and a.e. on $[\delta, T[\times \mathbb{R}$ to ρ . Finally, since $T > \delta > 0$ are arbitrary, the proof of (ii) is complete.

• **STEP 3.** *Strong convergence for initial datum in \mathbf{BV} .*

Let $\bar{\rho}$ satisfy (InBV). The result in Proposition 3.5 ensures that both $(\hat{\rho}^n)_{n \in \mathbb{N}}$ and $(\check{\rho}^n)_{n \in \mathbb{N}}$ are uniformly bounded in $\mathbf{L}^\infty([0, +\infty[; \mathbf{BV}(\mathbb{R}))$. Hence, we can repeat the proof of Proposition 3.7 (we omit the details) to obtain that

$$\int_0^L |\check{\rho}^n(t, z) - \check{\rho}^n(s, z)| dz \leq [\text{Lip}(v|_{[0, R]}) \text{TV}(\bar{\rho}) + (v_{\max} - v(R))] |t - s|,$$

for all $t, s \geq 0$. Therefore, Helly's theorem implies the desired compactness. Moreover, we can use Theorem 3.2 with w being the identity function on $[0, R]$, and obtain the desired compactness of $\hat{\rho}^n$. \diamond

We now prove that the two limits $\check{\rho}$ and ρ are related.

Proposition 3.10. *Let F be the cumulative distribution of ρ as defined in Proposition 3.2. Then*

$$\check{\rho}(t, F(t, x)) = \rho(t, x) \quad \text{for a.e. } (t, x) \text{ in } \text{spt}(\rho).$$

Proof. By definition (3.6) and Lemma 2.2, for any $\varphi \in \mathbf{C}_c^\infty([0, T] \times \mathbb{R})$ we have

$$\begin{aligned} \int_0^T \int_0^L \check{\rho}^n(t, z) \varphi(t, \hat{X}^n(t, z)) \, dz \, dt &= \int_0^T \int_0^L \hat{\rho}^n(t, \hat{X}^n(t, z)) \varphi(t, \hat{X}^n(t, z)) \, dz \, dt \\ &= \int_0^T \int_{\mathbb{R}} \hat{\rho}^n(t, x)^2 \varphi(t, x) \, dx \, dt. \end{aligned}$$

By extracting the a.e. convergent subsequence provided in Proposition 3.9 (and by extracting, if necessary, a further subsequence), we can send $n \rightarrow +\infty$ in the above identity and use the Lebesgue dominated convergence theorem (as the support of $\check{\rho}^n$ and $\hat{\rho}^n$ are uniformly bounded w.r.t. n) to get

$$\int_0^T \int_0^L \check{\rho}(t, z) \varphi(t, X(t, z)) \, dz \, dt = \int_0^T \int_{\mathbb{R}} \rho(t, x)^2 \varphi(t, x) \, dx \, dt.$$

By changing variable $z = F(t, x)$ in the first integral above, we get

$$\int_0^T \int_{\mathbb{R}} \check{\rho}(t, F(t, x)) \rho(t, x) \varphi(t, x) \, dx \, dt = \int_0^T \int_{\mathbb{R}} \rho(t, x)^2 \varphi(t, x) \, dx \, dt, \quad (3.14)$$

and this concludes the proof. \diamond

In the next proposition we prove that ρ is a weak solution in the sense of (2.2).

Proposition 3.11. *The limit function ρ defined in Proposition 3.3 is a weak solution in the sense of (2.2).*

Proof. Let $\varphi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$. By (2.16), (3.5) and (3.6), for all n we have

$$\begin{aligned} &\int_{\mathbb{R}_+} \int_0^L \left[v(\check{\rho}^n(t, z)) \varphi_x(t, \tilde{X}^n(t, z)) \right] \, dz \, dt \\ &= \sum_{i=0}^{N_n-1} \int_{\mathbb{R}_+} \int_{i\ell_n}^{(i+1)\ell_n} \left[v(y_i^n(t)) \varphi_x(t, x_i^n(t)) \right] \, dz \, dt \\ &= \sum_{i=0}^{N_n-1} \int_{\mathbb{R}_+} \int_{i\ell_n}^{(i+1)\ell_n} \left[\dot{x}_i^n(t) \varphi_x(t, x_i^n(t)) \right] \, dz \, dt \\ &= \sum_{i=0}^{N_n-1} \int_{\mathbb{R}_+} \int_{i\ell_n}^{(i+1)\ell_n} \left[\frac{d}{dt} \varphi(t, x_i^n(t)) - \varphi_t(t, x_i^n(t)) \right] \, dt \\ &= - \int_0^L \varphi(0, \tilde{X}^n(0, z)) \, dz - \int_{\mathbb{R}_+} \int_0^L \varphi_t(t, \tilde{X}^n(t, z)) \, dz \, dt. \end{aligned}$$

Since $(\tilde{X}^n)_{n \in \mathbb{N}}$ and $(\check{\rho}^n)_{n \in \mathbb{N}}$ converge strongly in $\mathbf{L}^1([0, T] \times [0, L]; \mathbb{R})$, and in view of Proposition 3.4, we get by sending $n \rightarrow +\infty$

$$\int_{\mathbb{R}_+} \int_0^L \left[\varphi_t(t, X(t, z)) + v(\check{\rho}(t, z)) \varphi_x(t, X(t, z)) \right] \, dz \, dt + \int_0^L \varphi(0, X_{\bar{\rho}}(z)) \, dz = 0.$$

We now apply the change of variable $x = X(t, z)$, see Lemma 2.2, and obtain

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[\rho(t, x) \varphi_t(t, x) + \rho(t, x) v(\check{\rho}(t, F(t, x))) \varphi_x(t, x) \right] \, dx \, dt + \int_{\mathbb{R}} \bar{\rho}(x) \varphi(0, x) \, dx = 0.$$

Finally, by Proposition 3.10 we have $\check{\rho}(t, F(t, x)) = \rho(t, x)$ a.e. on $\text{spt}(\rho)$, and therefore ρ satisfies (2.2). \diamond

We are now ready to complete the proof of our main result.

Proof of Theorem 2.3. In view of Theorem 2.2, the entropy inequality (2.4) is sufficient in order to show that ρ is the unique entropy solution in the sense of Definition 2.1.

Let $\varphi \in C_c^\infty([0, +\infty[\times \mathbb{R})$ with $\varphi \geq 0$ and $k \geq 0$ be a constant. We shall prove that the limit ρ satisfies the entropy inequality (2.4). We consider the quantity

$$\begin{aligned} & \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[|\hat{\rho}^n(t, x) - k| \varphi_t(t, x) + \operatorname{sgn}(\hat{\rho}^n(t, x) - k) [f(\hat{\rho}^n(t, x)) - f(k)] \varphi_x(t, x) \right] dx dt \\ &= B_0 + B_N + \sum_{i=0}^{N_n-1} I_i, \end{aligned}$$

with

$$\begin{aligned} B_0 &\doteq \int_{\mathbb{R}_+} \int_{-\infty}^{x_0^n(t)} \left[k \varphi_t(t, x) + f(k) \varphi_x(t, x) \right] dx dt, \\ B_N &\doteq \int_{\mathbb{R}_+} \int_{x_N^n(t)}^{+\infty} \left[k \varphi_t(t, x) + f(k) \varphi_x(t, x) \right] dx dt, \\ I_i &\doteq \int_{\mathbb{R}_+} \int_{x_i^n(t)}^{x_{i+1}^n(t)} |y_i^n(t) - k| \varphi_t(t, x) dx dt \\ &\quad + \int_{\mathbb{R}_+} \int_{x_i^n(t)}^{x_{i+1}^n(t)} \operatorname{sgn}(y_i^n(t, x) - k) [f(y_i^n(t, x)) - f(k)] \varphi_x(t, x) dx dt. \end{aligned}$$

For simplicity in the notation, from now on we shall drop the n index and the (t, x) dependency, except in cases in which $t = 0$. Moreover we define $y_N \equiv 0$. We next observe by (2.16) that

$$\frac{d}{dt} \left[\int_{x_i}^{x_{i+1}} \varphi dx \right] = v(y_{i+1}) \varphi(t, x_{i+1}) - v(y_i) \varphi(t, x_i) + \int_{x_i}^{x_{i+1}} \varphi_t dx, \quad (3.15)$$

$$\frac{d}{dt} \left[\int_{-\infty}^{x_0} \varphi dx \right] = v(y_0) \varphi(t, x_0) + \int_{-\infty}^{x_0} \varphi_t dx, \quad (3.16)$$

$$\frac{d}{dt} \left[\int_{x_N}^{+\infty} \varphi dx \right] = -v_{\max} \varphi(t, x_N) + \int_{x_N}^{+\infty} \varphi_t dx. \quad (3.17)$$

In view of (3.16) and (3.17), the terms B_0 and B_N can be rewritten as follows

$$B_0 = \int_{\mathbb{R}_+} k [v(k) - v(y_0)] \varphi(x_0) dt, \quad B_N = \int_{\mathbb{R}_+} k [v_{\max} - v(k)] \varphi(x_N) dt.$$

As for the term I_i , we have for $i = 0, \dots, N-1$

$$\begin{aligned} I_i &= \int_{\mathbb{R}_+} |y_i - k| \left\{ \frac{d}{dt} \left[\int_{x_i}^{x_{i+1}} \varphi dx \right] - v(y_{i+1}) \varphi(x_{i+1}) + v(y_i) \varphi(x_i) \right\} dt \\ &\quad + \int_{\mathbb{R}_+} \operatorname{sgn}(y_i - k) [f(y_i) - f(k)] [\varphi(x_{i+1}) - \varphi(x_i)] dt. \end{aligned}$$

By (2.18), we compute the term

$$\begin{aligned} & \int_{\mathbb{R}_+} |y_i - k| \frac{d}{dt} \left[\int_{x_i}^{x_{i+1}} \varphi dx \right] dt = - \int_{\mathbb{R}_+} \left[\int_{x_i}^{x_{i+1}} \varphi dx \right] \frac{d}{dt} |y_i - k| dt \\ &= - \int_{\mathbb{R}_+} \operatorname{sgn}(y_i - k) \left[-\frac{y_i^2}{\ell} [v(y_{i+1}) - v(y_i)] \right] \left[\int_{x_i}^{x_{i+1}} \varphi dx \right] dt \end{aligned}$$

$$= \int_{\mathbb{R}_+} \operatorname{sgn}(y_i - k) y_i [v(y_{i+1}) - v(y_i)] \left[\int_{x_i}^{x_{i+1}} \varphi \, dx \right] dt.$$

Hence, we have

$$\sum_{i=0}^{N-1} I_{i-1} = \sum_{i=1}^N \int_{\mathbb{R}_+} J_i \, dt + \sum_{i=1}^N \int_{\mathbb{R}_+} K_i \varphi(x_i) \, dt + \int_{\mathbb{R}_+} L \varphi(t, x_0) \, dt - \int_{\mathbb{R}_+} M \varphi(t, x_N) \, dt,$$

with

$$\begin{aligned} J_i &\doteq \operatorname{sgn}(y_{i-1} - k) y_{i-1} [v(y_i) - v(y_{i-1})] \left[\int_{x_{i-1}}^{x_i} \varphi \, dx - \varphi(x_i) \right], \\ K_i &\doteq \operatorname{sgn}(y_{i-1} - k) y_{i-1} [v(y_i) - v(y_{i-1})] + |y_i - k| v(y_i) \\ &\quad - \operatorname{sgn}(y_i - k) [f(y_i) - f(k)] - |y_{i-1} - k| v(y_i) + \operatorname{sgn}(y_{i-1} - k) [f(y_{i-1}) - f(k)], \\ L &\doteq |y_0 - k| v(y_0) - \operatorname{sgn}(y_0 - k) [f(y_0) - f(k)], \\ M &\doteq k v_{\max} - f(k). \end{aligned}$$

We observe that

$$B_N - \int_{\mathbb{R}_+} M \varphi(t, x_N) \, dt = 0.$$

We now compute L . If $k < y_0$, we have

$$L = k [v(k) - v(y_0)] \geq 0,$$

as v is non increasing. Therefore, for $k < y_0$

$$B_0 + \int_{\mathbb{R}_+} L \varphi(t, x_0) \, dx = 2 \int_{\mathbb{R}_+} k [v(k) - v(y_0)] \varphi(x_0) \, dt \geq 0.$$

Similarly, for $k \geq y_0$ we have

$$L = k [v(y_0) - v(k)] \geq 0,$$

which gives

$$B_0 + \int_{\mathbb{R}_+} L \varphi(t, x_0) \, dx = 0.$$

We now compute the term K_i for $i = 1, \dots, N$. After some easy manipulations, we get

$$K_i = k [v(k) - v(y_i)] \{ \operatorname{sgn}(y_i - k) - \operatorname{sgn}(y_{i-1} - k) \}.$$

We consider all the possible cases for k . If either $k < \min\{y_i, y_{i-1}\}$, or $k > \max\{y_i, y_{i-1}\}$, then we easily get $K_i = 0$. If $y_i \leq k \leq y_{i-1}$, then $K_i = 2k[v(y_i) - v(k)] \geq 0$ as v is non increasing. Finally, if $y_{i-1} \leq k \leq y_i$, then $K_i = 2k[v(k) - v(y_i)] \geq 0$. In all cases, we get $K_i \geq 0$ for all $i = 1, \dots, N$. Putting all the terms together, we get

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left[|\hat{\rho} - k| \varphi_t + \operatorname{sgn}(\hat{\rho} - k) [f(\hat{\rho}) - f(k)] \varphi_x \right] dx \, dt \geq \sum_{i=1}^N \int_{\mathbb{R}_+} J_i \, dt. \quad (3.18)$$

We now estimate the terms J_i . For some $\delta > 0$, assuming that the support of φ is contained in the strip $t \in [\delta, T]$, we have by Proposition 3.6

$$\left| \sum_{i=1}^N \int_{\mathbb{R}_+} J_i \, dt \right| = \left| \sum_{i=1}^N \int_{\mathbb{R}_+} \operatorname{sgn}(y_{i-1} - k) y_{i-1} [v(y_i) - v(y_{i-1})] \left[\int_{x_{i-1}}^{x_i} \varphi \, dx - \varphi(x_i) \right] dt \right|$$

$$\begin{aligned}
&\leq \int_{\mathbb{R}_+} \sum_{i=1}^N \left[\frac{y_{i-1}^2}{\ell} |v(y_i) - v(y_{i-1})| \int_{x_{i-1}}^{x_i} |\varphi(x) - \varphi(x_i)| dx \right] dt \\
&\leq \text{Lip}(\varphi) \int_{\delta}^T \sup_{i=1, \dots, N} \left[\frac{y_{i-1}^2 (x_i - x_{i-1})^2}{\ell} \right] \sum_{i=1}^N |v(y_i) - v(y_{i-1})| dt \\
&\leq \ell \text{Lip}(\varphi) T \sup_{t \geq \delta} \text{TV}[v(\hat{\rho}^n(t))] \leq \ell \text{Lip}(\varphi) T C_{\delta}.
\end{aligned}$$

As a consequence

$$\lim_{n \rightarrow +\infty} \sum_{i=1}^N \int_{\mathbb{R}_+} J_i dt = 0$$

and letting n go to infinity in (3.18) we obtain the entropy inequality (2.4). \diamond

3.5. Concluding remarks. We conclude this paper with the some technical remarks which help motivating our choices in the strategy of the proof at several stages in the paper.

- In the case of v such that $v' \leq -c < 0$, then the Oleinik-type estimate (3.12) gives a one sided estimate for $\hat{\rho}_x^n$ in the sense of distributions. Such an estimate can be then passed to the limit very easily, and one obtains an analogous estimate for the limit. In this way, one can check that the limit ρ is an entropy solution in much easier way than the above proof. In the general case of v' possibly degenerating, such a strategy fails. Indeed, surprisingly enough the Oleinik estimate one gets in the limit from (3.12) is *not* equivalent (in general) to the estimate (2.5). For this reason, we preferred getting the entropy condition in the Kruřkov sense rather than the one sided Lipschitz condition. This strategy allows in particular to get the entropy condition in the limit also in the case of v not satisfying (V3) and $\bar{\rho}$ satisfying (InBV).
- In the case of linear velocity v , e.g. $v(\rho) = v_{\max}(1 - \rho)$, the convergence to a weak solution (2.2) can be obtained without the need of the **BV** estimates, as the velocity term in (3.7) is linear. This is somehow intrinsic in using a Lagrangian description.
- In order to get continuity in time for the sequence $\hat{\rho}^n$, the most natural try would be getting \mathbf{L}^1 -continuity. Encouraged by the \mathbf{L}^1 time equi-continuity of $\check{\rho}^n$, we have attempted at proving such a property in many ways without success. This is the reason why we use the generalized Aubin-Lions lemma, which allows to take advantage of the Wasserstein equi-continuity of $\hat{\rho}^n$, and still get the same \mathbf{L}^1 -compactness in the end. The only drawback of this strategy is that we can't get any \mathbf{L}^1 time continuity for the limit.
- As pointed out in the introduction, the proposed Lagrangian approach has the advantage of providing a piecewise constant approximation with a non increasing number of jumps. The price to pay for such a simplification is that we lose the classical shock structure at a microscopic level. Indeed, as pointed out in [15, 46], the explicit solution to the FTL system even for simple Riemann-type initial conditions is not immediate. On the other hand, this aspect gives an added value to our result, as we show that shocks and rarefaction waves are still achieved in the macroscopic limit, despite not being easily detectable at the microscopic level.
- We finally recall that a symmetric construction can be set in the case of monotone increasing velocities. In that case, the suitable particle system should be recursively defined 'from the left towards the right', and therefore each particles adjusts its velocity according to the distance of the particle at its left hand side.

APPENDIX A. HEURISTIC DERIVATION OF THE FTL MODEL FROM THE LWR MODEL

In this appendix we formally provide our derivation from the scalar conservation law (2.1) of a discrete approximating model of the form (2.6).

Let ρ be an entropy solution of (2.1) in the sense of Definition 2.1. We assume for simplicity that ρ is compactly supported. The total mass of ρ in $] -\infty, x]$ at time $t \geq 0$ is

$$F(t, x) \doteq \int_{-\infty}^x \rho(t, y) dy. \quad (\text{A.1})$$

Clearly, F takes values in $[0, L]$, where $L \doteq \|\bar{\rho}\|_{\mathbf{L}^1(\mathbb{R}; [0, 1])}$, and for any fixed $t \geq 0$ the map $x \mapsto F(t, x)$ is continuous and non-decreasing, $F(t, -\infty) = 0$ and $F(t, +\infty) = L$. The result in the next proposition shows that (2.1a) is equivalent to requiring that the weak partial derivatives of F with respect to time and space commute in the sense of distributions.

Proposition A.1 ([16]). *The partial derivatives of F satisfy in the sense of distributions*

$$F_x = \rho, \quad F_t = -f(\rho). \quad (\text{A.2})$$

Proof. The first equality in (A.2) is obvious. For any test function $\psi \in \mathbf{C}_c^\infty([0, +\infty[\times \mathbb{R}; \mathbb{R})$ we have that by (2.2)

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}_+} F(t, x) \partial_t \psi_x(t, x) dt dx &= \int_{\mathbb{R}} \int_{\mathbb{R}_+} F(t, x) \partial_x \psi_t(t, x) dt dx \\ &= - \int_{\mathbb{R}} \int_{\mathbb{R}_+} \rho(t, x) \psi_t(t, x) dt dx = \int_{\mathbb{R}} \int_{\mathbb{R}_+} f(\rho(t, x)) \psi_x(t, x) dt dx. \end{aligned}$$

This shows that for any $t \geq 0$, the map $x \mapsto [F_t(t, x) + f(\rho(t, x))]$ is constant (as a distribution). Therefore there exists $c \in \mathbf{L}_{\text{loc}}^1([0, +\infty[; \mathbb{R})$ such that

$$\int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[F(t, x) \varphi_t(t, x) - f(\rho(t, x)) \varphi(t, x) + c(t) \varphi(t, x) \right] dt dx = 0.$$

Choose now, for any integer $k \in \mathbb{N}$,

$$\varphi(t, x) = \eta(t) \psi(x - k),$$

where $\eta \in \mathbf{C}_c^\infty([0, +\infty[; \mathbb{R})$ and $\psi \in \mathbf{C}_c^\infty(\mathbb{R}; [0, +\infty[)$ such that $\|\psi\|_{\mathbf{L}^1(\mathbb{R}; \mathbb{R})} = 1$. We get

$$\begin{aligned} 0 &= \int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[F(t, x) \dot{\eta}(t) - f(\rho(t, x)) \eta(t) + c(t) \eta(t) \right] \psi(x - k) dt dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[F(t, x + k) \dot{\eta}(t) - f(\rho(t, x + k)) \eta(t) + c(t) \eta(t) \right] \psi(x) dt dx. \end{aligned}$$

By Lebesgue dominated convergence theorem, we can send k to $+\infty$ and get

$$0 = \int_{\mathbb{R}} \int_{\mathbb{R}_+} \left[L \dot{\eta}(t) + c(t) \eta(t) \right] \psi(x) dt dx = \int_{\mathbb{R}_+} c(t) \eta(t) dt,$$

and the above expression on the right hand side can be easily made nonzero by suitably choosing η , unless $c(t) = 0$ for a.e. $t \geq 0$, which proves the assertion. \diamond

For any $t \geq 0$ the map $x \mapsto F(t, x)$ is strictly increasing on the intervals where the density $x \mapsto \rho(t, x)$ is not zero and otherwise it is constant. Therefore we can introduce $X \doteq \mathcal{X}[F]$, the pseudo-inverse of F . Now, assume for simplicity that $\rho(t, x) > 0$ for all $(t, x) \in \text{spt}(\rho) = \{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : a(t) \leq x \leq b(t)\}$. Then, for any $t \geq 0$ by Proposition A.1 we have that $x \mapsto F(t, x)$ is strictly increasing on $\text{spt}(\rho(t))$. This implies that $\mathcal{X}[F]$ is the inverse of F on the support of ρ , namely $F(t, X(t, z)) = z$ on $(t, z) \in \mathbb{R}_+ \times [0, L]$, and, assuming that all the derivatives below are well defined, we have that

$$F_x(t, X(t, z)) = \rho(t, X(t, z)) > 0 \quad \text{for a.e. } (t, z) \in \mathbb{R}_+ \times [0, L].$$

Therefore,

$$\begin{aligned} 1 &= \frac{d}{dz} F(t, X(t, z)) = F_x(t, X(t, z)) X_z(t, z), \\ 0 &= \frac{d}{dt} F(t, X(t, z)) = F_t(t, X(t, z)) + F_x(t, X(t, z)) X_t(t, z), \end{aligned}$$

which yields, once again by Proposition A.1, that $X(t, z)$ is indeed a solution of the PDE

$$X_t(t, z) = v \left(\frac{1}{X_z(t, z)} \right). \quad (\text{A.3})$$

The initial condition $X(0, z)$ is determined by

$$\int_{-\infty}^{X(0, z)} \bar{\rho}(y) dy = z.$$

The computation above is only rigorous on the sets in which $\rho(t, x) > 0$.

The last step needed in order to (formally) recognize the discrete model (2.6) in (A.3) is by replacing the z -derivative of X in (A.3) by the (forward) finite differences

$$X_z \approx \frac{X(t, z + \ell) - X(t, z)}{\ell}, \quad (\text{A.4})$$

which gives

$$X_t(t, z) \approx v \left(\frac{\ell}{X(t, z + \ell) - X(t, z)} \right).$$

Then the desired model (2.6) is obtained by assuming that $X(t)$ is piecewise constant on intervals of measure ℓ , with $X(t, j\ell) = x_j(t)$, $j = 1, \dots, N-1$. For any fixed $z \in \{i\ell : i = 0, 1, \dots, N\}$, the map $t \mapsto X(t, z)$ can be ideally interpreted as the path described by the ‘infinitesimal particle’ labelled with $z \in [0, L]$. Therefore, (A.3) can be interpreted as the expression in the Lagrangian coordinates (t, z) of the Cauchy problem (2.1).

ACKNOWLEDGEMENTS

Both the authors acknowledge useful discussions with J.A. Carrillo, R.M. Colombo, P. Degond, P. Marcati, and D. Matthes. M. Di Francesco is supported by the Marie Curie CIG (Career Integration Grant) DifNonLoc - Diffusive Partial Differential Equations with Nonlocal Interaction in Biology and Social Sciences and by the Ministerio de Ciencia e Innovación, grant MTM2011-27739-C04-02. M.D. Rosini is supported by ICM. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji nr: DEC-2011/01/B/ST1/03965.

REFERENCES

- [1] Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2005) 1.2
- [2] Aubin, J.P.: Macroscopic traffic models: Shifting from densities to ‘celerities’. Applied Mathematics and Computation **217**(3), 963 – 971 (2010) 1.3
- [3] Aw, A., Klar, A., Materne, T., Rascle, M.: Derivation of continuum traffic flow models from microscopic follow-the-leader models. SIAM Journal on Applied Mathematics **63**(1), 259–278 (2002) 1.5
- [4] Aw, A., Rascle, M.: Resurrection of ‘second order’ models of traffic flow. SIAM Journal on Applied Mathematics **60**(3), 916–938 (2000) 1.5
- [5] Bellomo, N., Bellouquid, A., Nieto, J., Soler, J.: On the multiscale modeling of vehicular traffic: from kinetic to hydrodynamics. Discrete and continuous dynamical systems B **19**, 1869–1888 (2014) 1.5
- [6] Berthelin, F., Degond, P., Delitala, M., Rascle, M.: A model for the formation and evolution of traffic jams. Arch. Ration. Mech. Anal. **187**(2), 185–220 (2008) 1.4, 1.5
- [7] Bolley, F., Brenier, Y., Loeper, G.: Contractive metrics for scalar conservation laws. J. Hyperbolic Differ. Equ. **2**(1), 91–107 (2005) 1.3

- [8] Brenier, Y., Grenier, E.: Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.* **35**(6), 2317–2328 (electronic) (1998) 1.2
- [9] Brenier, Y., Osher, S.: The discrete one-sided Lipschitz condition for convex scalar conservation laws. *SIAM J. Numer. Anal.* **25**(1), 8–23 (1988) 1.4
- [10] Bressan, A.: Global solutions of systems of conservation laws by wave-front tracking. *J. Math. Anal. Appl.* **170**(2), 414–432 (1992) 1.2
- [11] Bressan, A.: Hyperbolic systems of conservation laws, *Oxford Lecture Series in Mathematics and its Applications*, vol. 20. Oxford University Press, Oxford (2000) 1.1, 1.2, 2.1, 2.1, 3.1, 3.4
- [12] Carrillo, J.A., Di Francesco, M., Lattanzio, C.: Contractivity of Wasserstein metrics and asymptotic profiles for scalar conservation laws. *J. Differential Equations* **231**(2), 425–458 (2006) 1.3
- [13] Chen, G.Q., Rascle, M.: Initial layers and uniqueness of weak entropy solutions to hyperbolic conservation laws. *Arch. Ration. Mech. Anal.* **153**(3), 205–220 (2000) 2.1, 2.2
- [14] Colombo, R.M., Marson, A.: A Hölder continuous ODE related to traffic flow. *Proc. Roy. Soc. Edinburgh Sect. A* **133**(4), 759–772 (2003) 1.5
- [15] Colombo, R.M., Rossi, E.: On the micro-macro limit in traffic flow. To appear on *Rendiconti dell’Università di Padova* (2013) 1.2, 1.5, 3.5
- [16] Courant, R., Friedrichs, K.O.: *Supersonic Flow and Shock Waves*. Applied Mathematical Sciences. Springer-Verlag (1976) A.1
- [17] Dafermos, C.M.: Polygonal approximations of solutions of the initial value problem for a conservation law. *J. Math. Anal. Appl.* **38**, 33–41 (1972) 1.2, 2.1
- [18] Dafermos, C.M.: Hyperbolic conservation laws in continuum physics, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 325. Springer-Verlag, Berlin (2000) 1.1, 2.1, 2.1, 2.2
- [19] Daganzo, C.F.: A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological* **39**(2), 187–196 (2005) 1.3
- [20] Degond, P., Delitala, M.: Modelling and simulation of vehicular traffic jam formation. *Kinetic and related models* **1**, 279–293 (2008) 1.5
- [21] Di Francesco, M., Matthes, D.: Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. PDEs* DOI 10.1007/s00526-013-0633-5 (2013) 3.3
- [22] DiPerna, R.J.: Global existence of solutions to nonlinear hyperbolic systems of conservation laws. *J. Differential Equations* **20**(1), 187–212 (1976) 1.2
- [23] Dobrušin, R.L.: Vlasov equations. *Funktsional. Anal. i Prilozhen.* **13**(2), 48–58, 96 (1979) 1.2
- [24] Evans, L.C., Gariepy, R.F.: *Measure theory and fine properties of functions*, vol. 5. CRC press (1991) 3.3
- [25] Glimm, J.: Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.* **18**, 697–715 (1965) 1.2
- [26] Golse, F., Perthame, B.: Optimal regularizing effect for scalar conservation laws. *Rev. Mat. Iberoam.* **29**(4), 1477–1504 (2013). DOI 10.4171/RMI/765. URL <http://dx.doi.org/10.4171/RMI/765> 1.4
- [27] Goodman, J.B., LeVeque, R.J.: A geometric approach to high resolution TVD schemes. *SIAM J. Numer. Anal.* **25**(2), 268–284 (1988) 1.4
- [28] Greenberg, H.: An analysis of traffic flow. *Operations Research* **7**(1), 79–85 (1959) 2.4
- [29] Greenshields, B.: A study of traffic capacity. *Proceedings of the Highway Research Board* **14**, 448–477 (1935) 2.4
- [30] Hoff, D.: The Sharp Form of Oleinik’s Entropy Condition in Several Space Variables. *Transactions of the American Mathematical Society* **276**(2), 707–714 (1983) 1.4, 2.1
- [31] Hoogendoorn, S.P., Bovy, P.H.L.: State-of-the-art of vehicular traffic flow modelling. In: *Delft University of Technology, Delft, The*, pp. 283–303 (2001) 1.5
- [32] Kružkov, S.N.: First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)* **81 (123)**, 228–255 (1970) 1.1, 1.2, 1.6, 2.1, 2.1
- [33] Lax, P.D.: *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1973). Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 11 2.1
- [34] Lighthill, M.J., Whitham, G.B.: On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London. Ser. A.* **229**, 317–345 (1955) 1.5
- [35] MacCamy, R.C., Socolovsky, E.: A numerical procedure for the porous media equation. *Comput. Math. Appl.* **11**(1-3), 315–319 (1985). *Hyperbolic partial differential equations, II* 1.2
- [36] Matthes, D., Osberger, H.: Convergence of a variational lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.* **48**, 697–726 (2014) 1.2
- [37] Morrey Jr., C.B.: On the derivation of the equations of hydrodynamics from statistical mechanics. *Comm. Pure Appl. Math.* **8**, 279–326 (1955) 1.2
- [38] Newell, G.F.: A simplified theory of kinematic waves in highway traffic. *Transportation Research Part B: Methodological* **27**(4), 281–313 (1993) 1.3

- [39] Oleinik, O.A.: Discontinuous solutions of non-linear differential equations. *Uspekhi Matematicheskikh Nauk* **12**(3), 3–73 (1957) 1.1
- [40] Oleinik, O.A.: Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl. (2)* **26**, 95–172 (1963) 1.2, 1.4, 1.6, 2.1
- [41] Onsager, L.: Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)* **65**, 117–149 (1944) 1.2
- [42] Piccoli, B., Tosin, A.: Vehicular traffic: A review of continuum mathematical models. In: R.A. Meyers (ed.) *Encyclopedia of Complexity and Systems Science*. Springer New York (2009) 1.5
- [43] Pipes, L.A.: Car following models and the fundamental diagram of road traffic. *Transp. Res.* **1**, 21–29 (1967) 2.4
- [44] Richards, P.I.: Shock waves on the highway. *Operations Res.* **4**, 42–51 (1956) 1.5
- [45] Rosini, M.D.: Macroscopic models for vehicular flows and crowd dynamics: theory and applications. *Understanding Complex Systems*. Springer, Heidelberg (2013). DOI 10.1007/978-3-319-00155-5. URL <http://dx.doi.org/10.1007/978-3-319-00155-5> 1.5
- [46] Rossi, E.: On the micro-macro limit in traffic flow. Master’s thesis, Università Cattolica del Sacro Cuore, Brescia (2013) 1.5, 2.2, 3.5
- [47] Rossi, R., Savaré, G.: Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **2**(2), 395–431 (2003) 3.3, 3.1
- [48] Russo, G.: Deterministic diffusion of particles. *Comm. Pure Appl. Math.* **43**(6), 697–733 (1990) 1.2
- [49] Serre, D.: *Systems of conservation laws. 1 & 2*. Cambridge University Press, Cambridge (1999) 1.1, 2.1
- [50] Tadmor, E.: The large-time behavior of the scalar, genuinely nonlinear Lax-Friedrichs scheme. *Math. Comp.* **43**(168), 353–368 (1984) 1.4
- [51] Underwood, R.T.: Speed, volume, and density relationship. In: *Quality and theory of traffic flow: a symposium*, pp. 141–188. Greenshields, B.D. and Bureau of Highway Traffic, Yale University (1961) 2.4
- [52] Villani, C.: Topics in optimal transportation, *Graduate Studies in Mathematics*, vol. 58. American Mathematical Society, Providence, RI (2003) 1.2, 2.3, 2.3, 2.3
- [53] Zhang, H.M.: A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological* **36**(3), 275–290 (2002) 1.5

MARCO DI FRANCESCO - DEPARTMENT OF INFORMATION ENGINEERING, COMPUTER SCIENCE, AND MATHEMATICS - UNIVERSITY OF L’AQUILA, VIA VETOIO 1, 67100 L’AQUILA, ITALY.

E-mail address: mdifrance@gmail.com

MASSIMILIANO D. ROSINI - ICM, UNIVERSITY OF WARSAW, UL. PROSTA 69, 00-838 WARSAW, POLAND

E-mail address: mrosini@icm.edu.pl